



This work is protected by copyright and other intellectual property rights and duplication or sale of all or part is not permitted, except that material may be duplicated by you for research, private study, criticism/review or educational purposes. Electronic or print copies are for your own personal, non-commercial use and shall not be passed to any other individual. No quotation may be published without proper acknowledgement. For any other use, or to quote extensively from the work, permission must be obtained from the copyright holder/s.

SYNCHROTRON X-RADIATION CRYSTALLOGRAPHIC STUDIES
OF 6-PHOSPHOGLUCONATE DEHYDROGENASE FROM
BACILLUS STEAROTHERMOPHILUS

Thesis submitted in partial fulfilment for
the degree of
Doctor of Philosophy
at the
University of Keele
December 1984

Paul David Carr B.Sc. (Hons).

Department of Physics

DECLARATION

I hereby declare that the foregoing is the true and correct copy of the original as the same appears in my possession and control.

Paul E. Case
PAUL E. CASE
SACRAMENTO, CALIF.

To Mum and Dad
and Leslie

DECLARATION

I hereby declare that, unless otherwise in the text, the work presented herein is my own and has not been previously submitted for a degree at this or any other university.

Paul Carr

PAUL D. CARR

DECEMBER 1984

ACKNOWLEDGEMENTS

I would like firstly to thank my supervisor, Dr. John Helliwell for his guidance, encouragement and boundless enthusiasm at all stages of this work. I would like to thank Professor Watson Fuller for allowing me to work in his laboratory within its friendly and constructive atmosphere. The University of Keele Physics Department Workshop are thanked for their excellent technical support. A particular debt of gratitude is owed to Dr. Trevor Greenhough from whose vast crystallographic knowledge I drank greedily whilst at Keele.

The staff at Daresbury Laboratory are thanked for their friendship and assistance, in particular Dr. Mike Elder's applications group; Pella Machin and John Campbell for always being there and willing to sort out their many CCP programs; Colin Jackson and Sue Kennerly who scanned and rescanned my films without complaint.

Eleanor Dodson is warmly thanked for her guidance and hospitality during my stay at York. Drs. Margaret Adams and Richard Pickersgill from Oxford are thanked for supplying the sheep liver 6PGDH atomic coordinates.

Drs. Steve Rule, Miroslav Papiz, John Smith and Bob Liddington are thanked for their many useful discussions.

Thanks are also due to the UMIST Advanced Graphics Unit for allowing me to use their facilities in order to produce the diagrams contained within this thesis, in particular to Alan Goodkin who taught me how to use the system. Mrs.Clair Sacks is thanked for typing the manuscript.

I am pleased to acknowledge financial support from the University of Keele and thank the S.E.R.C. for allowing me to use the SRS and their computing services at Daresbury.

Finally, my biggest thanks of all go to my wife, Leslie, who has supported and encouraged me throughout and who has kept me sane this last year.

ABSTRACT

The synchrotron X-radiation crystallographic study of the enzyme 6-phosphogluconate dehydrogenase (6-phospho-D-gluconate NADP+ oxidoreductase (decarboxylating), EC 1.1.1.44) extracted from Bacillus stearothermophilus is reported. This crystallographic data was unobtainable using conventional sources. During this project I was involved in the building and commissioning of the central facility for protein crystallography at the SRS, Daresbury, (Helliwell et al (1982)). Chapter Three gives a detailed description of the instrument along with the relevant theory required for calculating beam compression, Guiner position and spectral resolution. Further expressions for optimising the beam geometry are derived from first principles. Chapter Four describes the data collection from bacterial 6PGDH crystals and details the data reduction procedures used to obtain a final dataset of integrated intensities using the MOSCO suite of programs. There is also a discussion of radiation damage in protein crystals within this chapter.

Chapter One gives background information on the bacterial 6PGDH enzyme including its preparation and properties. It also discusses thermal stability in microorganisms, concentrating particularly on proteins from the thermophilic bacteria. In addition, there is a

discussion of the relationships between the various dehydrogenase structures that have to date been solved. A summary of the relevant diffraction theory applicable to protein crystallography is given in Chapter Two.

The technique of molecular replacement was used in an attempt to obtain initial phases for the bacterial data. The theory of the method and the results of our calculations are given in Chapters Five and Six. The molecular replacement calculations revealed a direction of the non-crystallographic dimer axis and also allowed us to determine the correct space group as $P3_221$. The thesis is concluded with some closing remarks and suggestions for further work.

Symbols and Abbreviations

$\alpha\beta\gamma$	Eulerian angles
$\omega\phi\chi$	spherical polar coordinates
$\xi\zeta\tau$	cylindrical polar coordinates
$\underline{a}^*, \underline{b}^*, \underline{c}^*$	reciprocal lattice vectors
$h\ k\ l$	Miller indices
$x\ y\ z$	cartesian coordinates in real space
$x_f\ y_f$	film coordinates
$DC1, DC2, DC3$	directional cosines
$[\Omega]$	rotation matrix (o/p from rotation function)
$[\Phi]$	rotation matrix corresponding to spindle axis
$[A]$	rotation matrix defining orientation of reciprocal lattice vectors for a perfectly set crystal
$[\psi]$	misorientation matrix
\underline{F}	structure factor
\underline{F}^*	complex conjugate of \underline{F}
\underline{h}	reciprocal lattice vector
$G_{\underline{h}\underline{h}'}$	interference function
\underline{t}	translation vector
\underline{s}	scattering vector
$G(\underline{s})$	molecular transform
f	form factor
θ	Bragg angle
α	asymmetric angle of cut
r	$\theta - \alpha$
i	$\theta + \alpha$

λ	wavelength
ϕ	spindle angle
ω_{acc}	rocking width of incident beam
β	convergence angle of incident beam
θ_x	angle of crossfire
Δ	reflecting range of equatorial reflections
ϵ	radius of reciprocal lattice point
η	crystal mosaicity
μ	absorption coefficient
ρ	electron density
p	source to monochromator distance
p'	monochromator to focus distance
$\delta\lambda$	spectral resolution
λ	
m	index of partiality
k	scale factor
B	temperature factor
k_e, k_b	absorption coefficients of film emulsion and base
L_p^{-1}	Lorentz - polarisation factor
hsw	hand-switch option
e	charge on an electron
m	mass of an electron
c	speed of light
I	intensity
e.m	electromagnetic
eV	electron-volt
K_m	Michaelis constant

°	°
Å	Ångstrom unit (10^{-10}m)
R	reliability index
RF	rotation function
SRS	Synchrotron Radiation Source, Daresbury, U.K.
SR	synchrotron radiation
SERC	Science & Engineering Council of Great Britain
EMBL	European Molecular Biology Laboratory
EXAFS	Extended X-ray Absorption Fine Structure
6PGDH	6-phosphogluconate dehydrogenase
6PG	6 phosphogluconate
LADH	liver alcohol dehydrogenase
MIR	Multiple Isomorphous Replacement
R-5-P	ribulose-5-phosphate
NADP	nicotinamide adenine dinucleotide phosphate
LDH	Lactate dehydrogenase
MDH	Malate dehydrogenase
GAPDH	glyceraldehyde phosphate dehydrogenase
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
EDTA	ethylene diamine tetra-acetic acid
β -Me	β -mercaptoethane
DHFR	dihydrofolate reductase
GR	Glutathione reductase
PHBH	p-hydroxybenzoate hydroxylase
FAD	Flavin adenine dinucleotide

All other symbols are as defined, in context, in the text.

CONTENTS

Page

CHAPTER ONE - INTRODUCTION

Introduction	1
1.1 The Pentose Phosphate Pathway	2
1.2 Preparation of the Enzyme from <u>Bacillus</u> <u>Stearothermophilus</u>	2
1.3 Crystallization	4
1.4 Properties of the Enzyme	5
1.4.1 Kinetic parameters	5
1.4.2 Inhibition and activation studies	5
1.4.3 Physical properties	6
1.4.4 Stability	7
1.5 Thermal Stability in Microorganisms	9
1.5.1 Proteins from thermophilic bacteria	10
1.6 Reasons for Studying the Enzyme	15
1.7 Relationship Between Dehydrogenase Structures	16
1.8 Thesis Rationale	19

CHAPTER TWO - CRYSTALLOGRAPHIC THEORY

2.1 Diffraction Theory	21
2.2 The Atomic Scattering Factor	23
2.3 The Molecular Transform	25

2.4	Structure Factor Equation	26
2.5	Reciprocal Lattice and Ewald's Construction	30
2.6	The Electron Density Equation	34
2.7	The Phase Problem	35
2.8	The Rotation Method	37

CHAPTER THREE - EXPERIMENTAL SET-UP AND DATA COLLECTION

3.1	Introduction	39
3.2	Synchrotron Radiation	40
3.3	The Mirror	43
3.4	Monochromatisation of the X-Ray Beam	44
3.4.1	The Guinier position	45
3.4.2	Spectral resolution	47
3.4.3	Horizontal spot width	48
3.4.4	Fraction of the available synchrotron beam transmitted via the collimator	50
3.4.5	Depth of focus	53
3.5	The Experimental Workstation	53
3.5.1	Slit design	55
3.5.2	The monochromator vessel	56
3.5.3	The experimental arm	56
3.5.4	The X-ray camera and alignment carriage	60
3.5.5	The computer control system	60
3.5.6	Alignment procedure	63
3.5.7	Radiation protection	65
3.6	Conclusion	67

CHAPTER FOUR - DATA ACQUISITION AND REDUCTION

4.1	Data Collection	69
4.2	Data Processing	70
4.2.1	Refinement of crystal orientation and unit cell parameter	71
4.2.2	Prediction of reflections and their partiality	73
4.2.3	Recording of integrated intensities and background measurements	76
4.2.4	Post-refinement	80
4.2.5	Correction of integrated intensities for geometrical factors	81
4.2.6	Film to film scaling	83
4.2.7	Crystal to crystal scaling	84
4.2.8	Radiation damage	85

CHAPTER FIVE - THE ROTATION FUNCTION

5.1	Introduction	92
5.2	Rotation Function Theory	93
5.3	Bacterial 6PGDH Self-Rotations	95
5.4	Bacterial/Sheep Liver 6PGDH Cross Rotation Functions	98
5.4.1	Cross rotations using new coordinates	106
5.4.2	Summary	108

CHAPTER SIX - THE TRANSLATION PROBLEM

6.1	Introduction	111
6.2	Method	112
6.3	Results	113
6.3.1	Results for the $P3_121$ cell	114
6.3.2	Results for the $P3_221$ cell	116

CHAPTER SEVEN - CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

7.1	Concluding Remarks	117
7.2	Suggestions for Further Work	121

REFERENCES	124
------------	-----

APPENDICES

CHAPTER 1

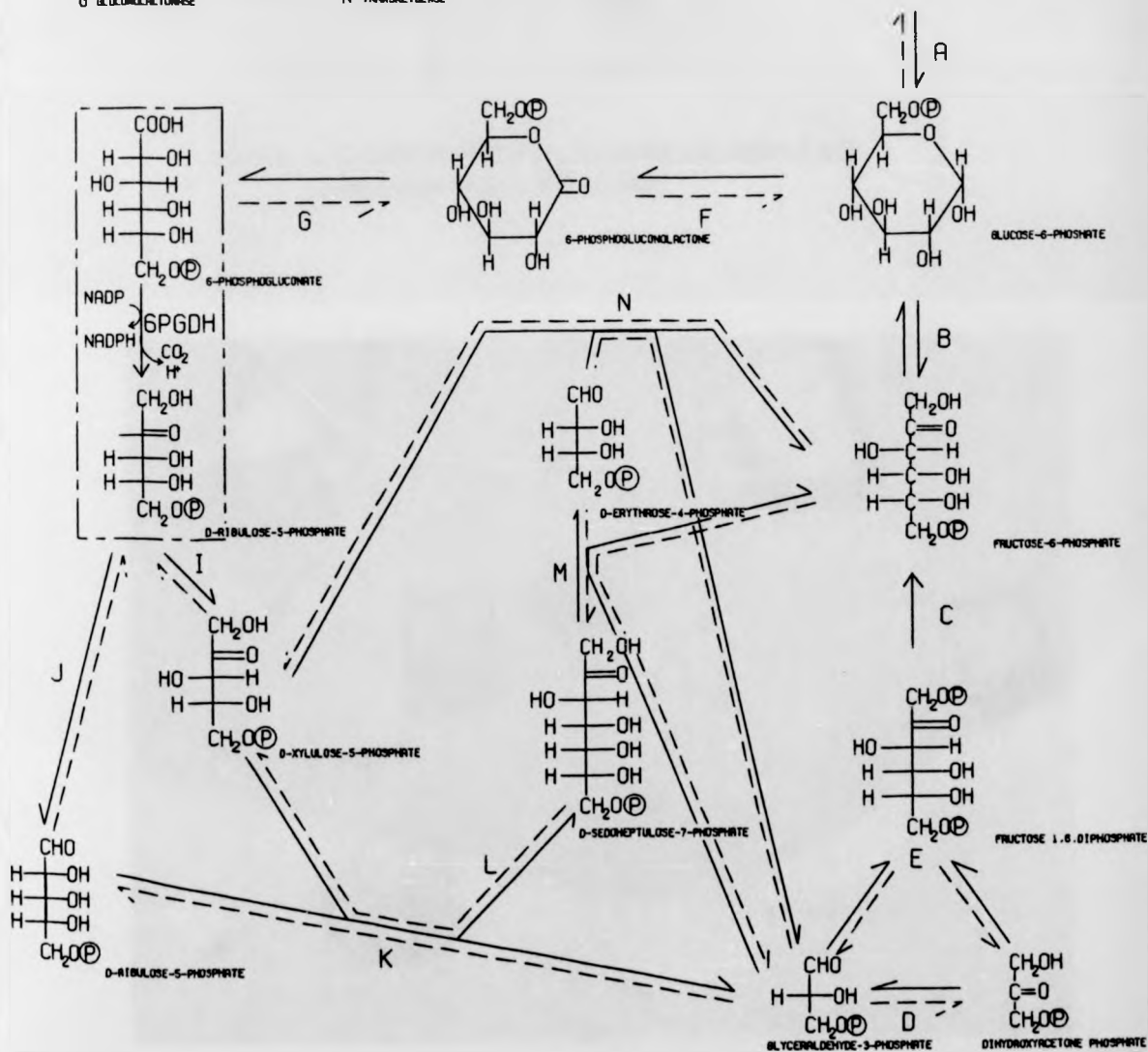
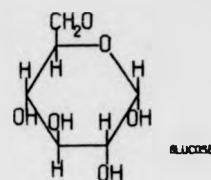
INTRODUCTION

6-phosphogluconate dehydrogenase (6-phospho-D-gluconate NADP⁺ oxidoreductase (decarboxylating), EC 1.1.1.44) is the third enzyme in the pentose pathway (see figure 1.1). 6 PGDH catalyses the conversion of 6-phosphogluconate (6PG) to ribulose-5-phosphate (R-5-P) and CO₂ utilizing the coenzyme nicotinamide adenine dinucleotide phosphate (NADP⁺ oxidised, NADPH reduced form). Two orientations of the hydrogen of the NADPH transferred in the oxidation-reduction reaction are possible and commonly called H_A (or H(R)) and H_B (or H(S)) orientations (see figure 1.2). Different dehydrogenases are either A or B side specific. For instance, LDH, LADH and MDH, which all utilise NAD⁺, are H_A specific whereas GAPDH and 6PGDH are H_B specific (Helliwell (1977)). The enzyme extracted from Bacillus stearothermophilus is of particular interest due to its ability to remain active at relatively high temperatures (see section 1.4.3).

Other oxidative decarboxylating enzymes do exist; malic enzyme catalysing the conversion of s-malate to pyruvate and CO₂; isocitrate dehydrogenase catalysing the conversion of 2S-3R-isocitrate to α -keto-glutarate and CO₂.

FIGURE 1.1 GLYCOLYTIC AND PENTOSE PATHWAYS

- | | |
|-------------------------------------|------------------------------------|
| A HEXOKINASE GLUCOSE-6-PHOSPHATE | H 6-PHOSPHOGLUCONATE DEHYDROGENASE |
| B PHOSPHOGLUCOSOMERASE | I PHOSPHOPENTOSE EPIMERASE |
| C FRUCTOSE-6-PHOSPHATASE | J PHOSPHOPENTOSE ISOMERASE |
| D TRIOS PHOSPHATE ISOMERASE | K PHOSPHOGLUCOSOMERASE |
| E ALDOLASE | L TRANSKETOLASE |
| F GLUCOSE-6-PHOSPHATE DEHYDROGENASE | M TRANSALDOLASE |
| G GLUCONOLACTONASE | N TRANSKETOLASE |



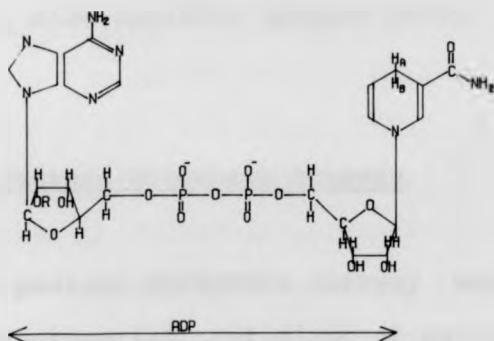


FIGURE 1.2 THE COENZYME, NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD).
WHEN R-PHOSPHATE IT BECOMES NADP

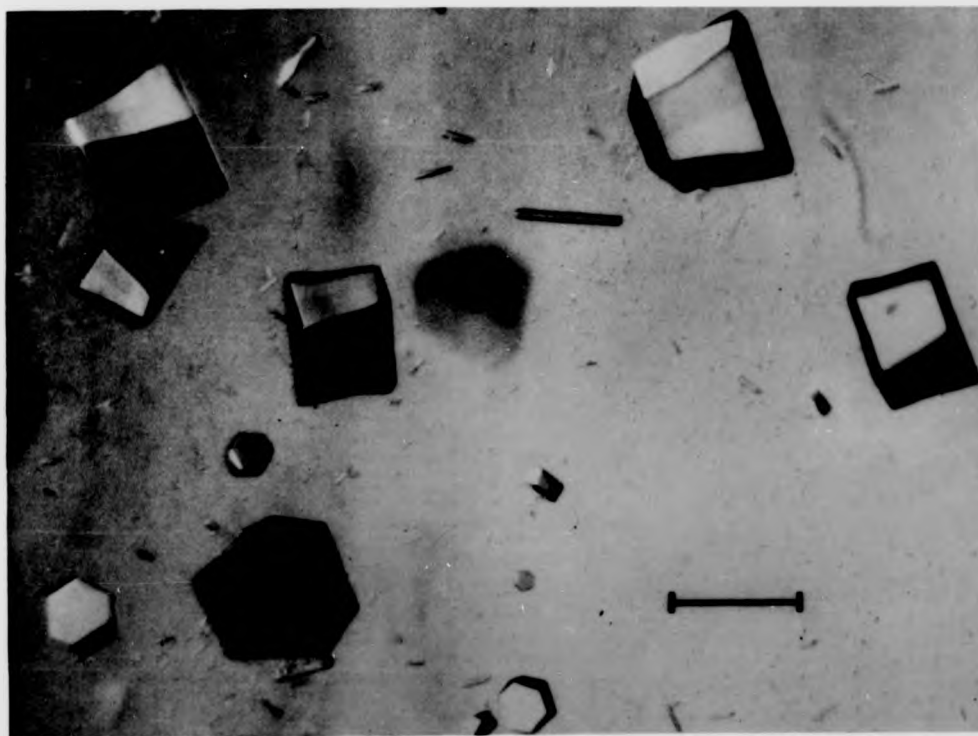


FIGURE 1.3 PHOTOGRAPH OF 6PGDH CRYSTALS FROM BACILLUS STEAROTHERMOPHILUS
THIS PHOTOGRAPH WAS KINDLY SUPPLIED BY DR. B.M.F. PEARSE TO WHOM I AM INDEBTED

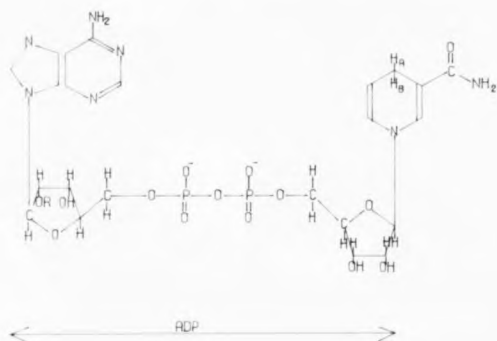


FIGURE 1.2 THE COENZYME, NICOTINAMIDE ADENINE DINUCLEOTIDE (NAD),
WHEN R-PHOSPHATE IT BECOMES NADP

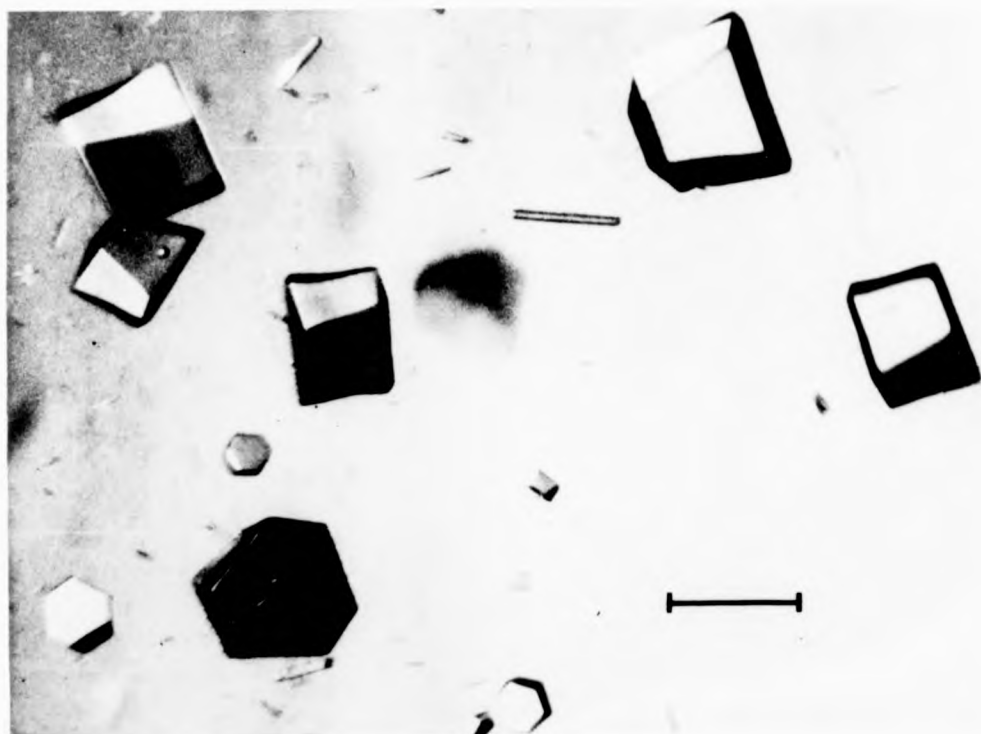


FIGURE 1.3 PHOTOGRAPH OF 6PGDH CRYSTALS FROM *BACILLUS STEAROTHERMOPHILUS*
THIS PHOTOGRAPH WAS KINDLY SUPPLIED BY DR. B.M.F. PEARSE TO WHOM I AM INDEBTED

Both the latter are, like 6PGDH, NADP⁺ dependent but unlike 6PGDH, H_A side specific (Popjak 1970).

1.1 The Pentose Phosphate Pathway

The pentose phosphate pathway occurs in the soluble extra mitochondrial cytoplasm of most organisms except very simple bacteria. The pathway has three major roles:

- (1) A source of energy.
- (2) The formation of pentose sugars, in particular ribose-5-phosphate, for incorporation into nucleotides and coenzymes (e.g. DNA, RNA, NAD and NADP).
- (3) The generation of NADPH (used in reductive synthesis, e.g. fatty acids, mononucleotides, amino acids and urea.)

The pathway is most important in the liver, mammary glands, adipose tissue and adrenal cortex of higher organisms.

1.2 Preparation of the Enzyme from Bacillus Stearothermophilus

A summary is now given (along with section 1.3) of the work done by Pearse and Harris (1973) which led to the

crystals used in the present study. Bacillus stearothermophilus (strain 1503) was grown at the Microbiological Establishment, Porton, Wilts, UK. Twenty kilograms of the bacterial cells were disrupted and the soluble proteins extracted, giving a solution containing 80,000 units of 6PGDH activity (1 unit of enzyme activity is defined as the number of μ moles of NADP reduced per minute at 25°C when assayed in 3ml tris-HCl buffer pH 9.0, I = 0.1 containing 0.1 mM NADP and 1.0 mM 6PGDH (Pearse and Harris (1973))).

The initial steps for the large scale purification of several enzymes were carried out at Porton (Atkinson et al (1972)). Fractionations were obtained using hydroxyapatite and DAEA-Sephadex. 16,000 units of 6PGDH activity, representing a 20% yield of enzyme, with a specific activity of about 1 unit/mg were present after separation from triose phosphate isomerase, aldolase and GAPDH. Subsequent steps leading to the isolation of the pure enzyme and initial crystallisation were carried out by Pearse and Harris (1973) at the MRC Laboratory of Molecular Biology, Cambridge, UK. The enzyme solution was subjected to ion-exchange chromatography on CM - Sephadex at pH 5.6, resulting in a 10-fold purification. The 6PGDH fraction was concentrated (in 0.1 I sodium acetate buffer pH 6.0 containing 1mM ethylene diamine tetra-acetic acid (EDTA), 0.1% β -mercaptoethanol (β -ME) and 10 μ M nicotinamide adenine dinucleotide phosphate (NADP) to give a total protein

concentration of about 20 mg/ml, with 200 units/ml of 6PGDH activity. Microcrystals of the enzyme formed, on standing, at 0-4°C. These were redissolved in 0.1 I acetate buffer pH 6.0. The enzyme (with a specific activity of 20 units/mg) was stored in this solution at 0-4°C, with a crystal of thymol to prevent bacterial growth.

The criterion used to determine the purity of the enzyme was that a redissolved crystal preparation gave a single band when examined by SDS-polyacrylamide gel electrophoresis. Electrophoresis of the protein on Cellogel at pH 8.6, with 10 μ M NADP in the electrode buffer gave a single band staining for protein and for 6PGDH activity. The homogeneity of the protein was also demonstrated by sedimentation velocity experiments at pH 8.0 and 0.1 I, a single symmetrical boundary with an $S_{20,w}$ value of 5.8 S was observed.

Characterisation of the enzyme and enzyme inhibition studies were also undertaken at Cambridge and are reported by Pearse and Harris (1973). Veronese et al (1982) report a similar procedure for preparation and also some further characterisation of 6PGDH from Bacillus stearothermophilus.

1.3 Crystallization

One ml of pure enzyme solution (4 mg/ml) was dialysed

against 200 ml of 0.1 I sodium acetate buffer pH 6.0, containing 1 mM EDTA, 0.1% (v/v) β -ME and 10 mM NADP, brought to 30% saturation with saturated ammonium sulphate solution. The protein was clarified by centrifuging at 10,000 revs/minute. Single μ l drops were placed in shallow depressions in a glass tray, which was then placed in a dish containing 6 ml of acetate buffer and 9 ml of saturated ammonium sulphate. The container was sealed and left undisturbed at room temperature. Crystals appeared after a few days. A photograph of the crystals is shown in figure 1.3.

1.4 Properties of the Enzyme

1.4.1 Kinetic Parameters

The enzyme is active in the pH range 6.0 to 9.0 and exhibits maximum activity around pH 8.0 both at 25°C and 43°C (Veronese et al (1974)). This is similar to that found for the enzyme extracted from Candida utilis (Rippa et al (1970)). The enzyme follows Michaelis-Menten kinetics in the range 25°C to 60°C and the apparent K_m values for NADP and 6-phosphogluconate at 43°C were found to be $2.5 \times 10^{-5}M$ and $2.0 \times 10^{-5}M$ respectively (Veronese et al (1974), (1982)). For a comparison of these K_m values with those obtained from 6PGDH extracted from other sources see table 1.1.

Table 1.1 Summary of some physical and kinetic properties of pure 6-phosphogluconate dehydrogenase from different sources. Table adapted from Helliwell (1977).

Source of enzyme	M.W.	No. Subunits	K _m (NADP) μM	K _m (6PG) μM	Ref
<u>B. stearothermophilus</u>	100,000	2	-	-	1
	100,000	2	25	20	2
<u>C. utilis</u>	101,000	2	20	52	3
<u>Neurospora crassa</u>	110,000- 120,000	2	30	10	4
<u>Streptococcus faecalis</u>	108,000 ± 3,600	2	15	24	5
rat liver	102,000 ± 3,000	2	13	71	6
sheep liver	94,000 ± 2,000	2	-	-	7
sheep liver	-	-	6.8	6.9	8
sheep liver	-	-	7.0	15	9
human erythrocyte	104,000	2	30	20	10
rabbit mammary gland	104,000	2	23	54	11

References

1. Pearse and Harris (1973).
2. Veronese, Boccu and Fontana (1982).
3. Rippa, Signorini and Pontremoli (1967).
4. Scott and Abramsky (1973).
5. Bridges, Palumbo and Wittenberger (1975).
6. Procsal and Holten (1972).
7. Silverberg and Dalziel (1973).
8. Villet and Dalziel (1972).
9. Dyson, D'Orazio and Hanson (1973).
10. Pearse and Rosemeyer (1974).
11. Betts and Meyer (1975).

1.4.2 Inhibition and Activation Studies

The enzyme from B.stearothermophilus, in common with 6PGDH from other sources shows sensitivity towards reagents that react with the sulphydryl (-SH) groups of proteins. The native enzyme was inactivated upon incubation in the presence of iodoacetamide, p-chloromercuribenzoate, Hg^{2+} ions, 5,5'-dithiobis-(2-nitrobenzoic acid) (Veronese et al (1974)) or 7-chloro-4-nitrobenzo-2-oxa-1,3-diazole (Fontana et al (1977)). Partial protection from these inhibitors could be gained from binding either substrate or NADP, implying that these reagents react in the vicinity of the active site (Veronese et al (1982)). With the exception of iodoacetamide, recovery from inactivation by the above can be obtained by adding an excess of β -ME (Fontana et al (1977)). Pearse and Harris (1973) report that radioactive mapping with $[\text{I}-^{14}\text{C}]$ iodoacetamide infer that two cystine molecules per dimer react with the inhibitor.

Veronese et al (1974) report that Mg^{2+} ions activate the enzyme at low concentrations but act as an inhibitor at high concentrations. They also show there to be an increase in the activation effect of times four when the temperature is raised from 30°C to 55°C .

1.4.3 Physical Properties

The molecular weight was determined from sucrose

density gradient centrifugation as 101,000 by Veronese et al (1974) using Fructose-1,6-diphosphate aldolase as an internal reference marker. They later modify this to be "around 100,000" (Veronese et al (1976),(1982)). Pearse and Harris (1973) obtained a value of 100,000 using sedimentation velocity experiments. They also showed it to be dimeric consisting of equal subunits of 50,000 by means of sodium dodecylsulphate (SDS) gel electrophoresis using bovine serum albumin, γ -globulin, GPDH and aldolase as marker proteins. Polyacrylamide gel electrophoresis also gave a single sharp band with a mobility corresponding to 50,000 daltons. These values are similar to those obtained for 6PGDH from other sources; for a comparison see table 1.1. The amino acid composition of the enzyme extracted from B.stearothermophilus according to Pearse (cited in Helliwell (1977)) and Veronese et al (1976) are given in table 1.2 along with the composition of 6PGDH from five other sources.

1.4.4 Stability

6PGDH from B.stearothermophilus retains full activity for several months when stored at 4°C as an ammonium sulphate suspension or in acetate buffer pH 6.0 (Pearse and Harris (1973)). Partially inactivated enzyme can be fully reactivated by incubation at 37°C in the presence of 1mM EDTA and β -ME (Veronese et al (1974)). The enzyme is stable to relatively high temperatures as one would expect from a

Table 1.2 Amino acid composition of 6PGDH from various sources.
Table adapted from Helliwell (1977).

Amino acid	<u>B.steoro- thermophilus</u>	<u>B.steoro- thermophilus</u> ²	sheep liver ³	<u>C. utilis</u> ⁴	Human red blood cell ⁵	<u>S. faecalis</u> ⁶
lysine	31	33	32	31	38	35
histidine	9	10	10	6	9	6
arginine	21	21	18	15	21	19
aspartate	37	39	41	46	50	43
threonine	19	18	16	19	20	25
serine	16	17	20	21	26	26
glutamate	57	53	33	45	45	69
proline	18	18	16	18	14	18
glycine	43	40	44	47	48	43
alanine	53	48	37	50	43	47
valine	25	28	21	26	29	17
methionine	11	10	11	8	12	13
isoleucine	40	33	30	15	28	33
leucine	38	38	41	38	42	40
tyrosine	19	19	11	18	12	22
phenylalanine	18	18	20	20	24	18
tryptophan	5	5	7	6	8	9
half-cystine	3	3	9	4	11	3
Total	463	451	417	433	480	486

References

1. B.M.F. Pearse (personal communication).
2. Veronese, Boccu and Fontana (1982).
3. Silverberg (1973).
4. Rippa, Signorini and Pontremoli (1972).
5. B.M.F. Pearse (Ph.D. thesis, University of London).
6. Bridges, Palumbo and Wittenberger (1975).

Table 1.3 Relationships between dehydrogenase structures. Table based on Adams et al (1981).

Enzyme	EC no.	Submit M.W.	Subait Structure	Coenzyme
Lactate dehydrogenase	1.1.1.27	35,000	Tetramer	NAD
α-Malate dehydrogenase	1.1.1.37	36,000	Dimer	NAD
Glyceraldehyde phosphate dehydrogenase	1.2.1.1	36,000	Tetramer	NAD
Liver alcohol dehydrogenase	1.1.1.1	40,000	Dimer	NAD
Dihydrofolate reductase	1.5.1.3	20,000	Monomer	NADP
Glutathione reductase	1.6.4.2	50,000	Dimer	NADP, FAD
p-Hydroxybenzoate hydroxylase	1.14.13.2	43,000	(Dimer)	NADP, FAD
6-Phosphogluconate dehydrogenase	1.1.1.44	47,000	Dimer	NADP

thermophile. 6PGDH from B.stearothermophilus is stable for at least two hours at a temperature of 60°C whereas the same enzyme extracted from E.coli is denatured at a temperature of 45°C. Figure 1.4 compares the heat curves of 6PGDH from the two bacterial sources. Figure 1.5 shows the heat inactivation curves for the Bacillus enzyme at a range of temperatures between 60°C and 80°C. Veronese (et al (1982)) correlated the thermal stability of enzymic activity with the secondary structure of the proteins by means of far-ultraviolet circular dichroism (CD) and fluorescence emission measurements. They also state that, as measured by these spectroscopic methods, the three dimensional structures of these two bacterial enzymes "do not possess particular structural features responsible for their enhanced stability; on the other hand, stability could arise from very subtle structural features".

In summary, the properties of the enzyme extracted from B. stearothermophilus are very similar to those from other sources of 6PGDH in terms of enzyme kinetics, molecular weight, amino acid composition and inactivation by various agents. On the other hand, the thermal stability and resistance to denaturants such as urea are far higher than in its mesophilic counterparts. The structural features which presumably afford this extra protection are not of a gross nature as shown by the spectroscopic results (Veronese et al (1974)).

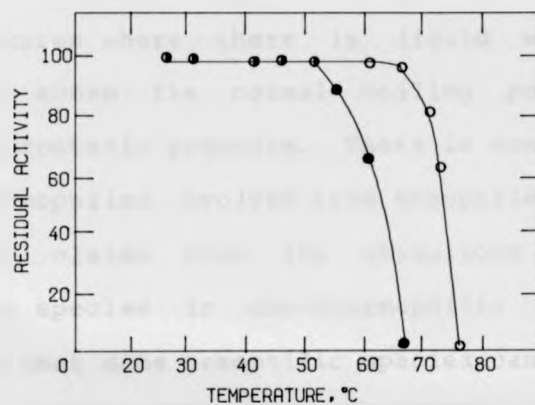


FIGURE 1.4 THE EFFECT OF HEAT ON THE ACTIVITY OF *BACILLUS STEAROTHERMOPHILUS* (○) AND *ESCHERICHIA COLI* (●) 6PGDH

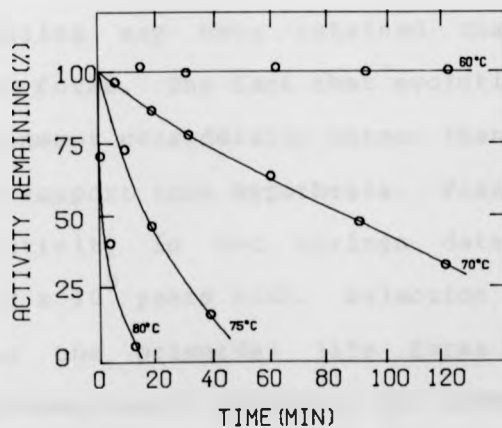


FIGURE 1.5 HEAT INACTIVATION OF 6PGDH FROM *BACILLUS STEAROTHERMOPHILUS*

1.5 Thermal Stability in Microorganisms

In a review of life at high temperature, Brock (1967) showed that thermophilic bacteria can be found growing at all temperatures where there is liquid water, including temperatures above the normal boiling point in areas of increased hydrostatic pressure. There is some debate as to whether thermophiles evolved from mesophiles or vice versa. Allen (1953) claims that the ubiquitous occurrence of thermophilic species in non-thermophilic environments and the finding that some mesophilic species can be adapted to grow at elevated temperatures is evidence that thermophiles have evolved from mesophiles. More recently, Tanaka et al (1971) have proposed phylogenetic trees of *Clostridium* sp. which indicate that thermophilic species are the oldest from evolutionary considerations. Brock (1967) has suggested that thermophiles may have retained characteristics of primordial life forms. The fact that evolution has proceeded from an environment considerably warmer than the present one would seem to support this hypothesis. Fossil evidence of bacterial activity in hot springs dates back to the Precambrian (2×10^9 years old). Selection is proposed to have acted on the primordial life forms by allowing more flexible conformations of proteins, for example, to exhibit greater substrate affinity by means of induced fit and allosteric mechanisms. Such mutants will have sacrificed their thermal stability but will have gained a selective advantage at lower temperatures. They could therefore grow

and function nearly as fast as their thermophilic counterparts and occupy a separate ecological niche.

There have been several mechanisms proposed to explain the survival of thermophilic microorganisms. They fall generally into three main categories; (i) Stabilization may be achieved through lipid interaction (including cellular membranes); (ii) Heat denatured cellular components may rapidly resynthesize; and (iii) thermophilic organisms possess macromolecular complexes with an inherent heat stability. Singleton and Amelunxen (1973) review the evidence for these mechanisms and conclude:

"Most of the available evidence supports the hypothesis that thermophilic microorganisms synthesize macromolecular components which possess an intrinsic thermostability..... one of the primary contributing factors toward the survival of the organism is the inherent heat stability of its cellular proteins. Without the presence of thermally stable biosynthetic systems, the organism could not survive."

1.5.1 Proteins from Thermophilic Bacteria

In general, protein molecules from thermophiles bear striking resemblance to their mesophilic counterparts. (Zuber (1975), Singleton and Amelunxen (1973)). Their points of similarity include:

- (1) Molecular weights;
- (2) Sub-unit composition;
- (3) Allosteric effectors;
- (4) Amino acid composition;
- (5) Primary sequences.

It is because of this great similarity in properties that it is important to obtain precise three dimensional structural information at atomic level via X-ray crystallographic techniques if we are to understand the subtle changes which give rise to the phenomenon of thermophily. By way of example, the substitution of a single residue (tyrosine 88 is replaced by histidine) in lysozyme from bacteriophage T4 (Matthews et al (1974)) causes a dramatic change in its thermal sensitivity. At 37°C, the activity of the mutant enzyme (histidine 88) is about 40% of that of the wild type (tyrosine 88); however, at 20°C, the activity is less than 0.01% of that of the wild type.

Despite the often subtle structural changes associated with thermophily, several similarities have been observed. In general, thermophilic proteins contain fewer free cysteine residues than their mesophilic counterparts, with the remaining ones often being used to form disulphide bonds. Examples where this difference is observed include Amylase (Manning et al (1961)) and Glyceraldehyde-3-phosphate dehydrogenase (Singleton et al (1969)). Harris (1975) points out that free cysteine residues are very likely to

have a destabilizing effect due to their susceptibility to oxidation and disulphide exchange reactions. It has also been observed that the formation of additional stabilising salt-bridges often occurs in thermophilic proteins (e.g. Ferredoxin and Haemoglobin (Perutz and Raidt (1975)), Glyceraldehyde 3-phosphate dehydrogenase (Biesecker et al (1977))). Metal ions have also been shown to greatly affect thermostability. Pfeuffer and Elliot (1969) observed that Ca^{2+} ions were required to give maximum thermal stability in Amylase and Levy et al (1972) have demonstrated that the binding of Mg^{2+} ions to yeast tRNA specific for phenylalanine caused a marked increase in thermostability.

A great deal of discussion has taken place regarding the role of hydrophobic residues in the stabilizing of thermophilic proteins. Koffler and Harbrough (1972) report that the thermal stability of Flagellin is effected by nitration of four of the six tyrosine residues present. The stability of an extracellular protease has been shown by Ohta et al (1966) to be enhanced by the presence of tyrosine and tryptophan residues existing in a hydrophobic environment. Further stabilisation was envisioned as arising from hydrogen bonding of 'abnormally titrating' tyrosine residues buried deep in a hydrophobic cluster. Scheraga (1963) and Brandts (1967) have pointed out that, unlike other bonding interactions, the strength of hydrophobic interactions increases with temperature, up to around 60°C.

Several authors (Howard and Becker (1970), Ljungdahl et al (1970) and Singleton et al (1969)) have calculated parameters which purportedly measure the capacity for hydrophobic interaction in a protein from its amino acid composition. These parameters are ϕH_{ave} (Bigelow (1967)), NPS (Waugh (1954)) and p (Fisher (1964)). The symbol ϕH_{ave} represents the average free energy change which would occur in transferring the amino acid residues of a protein from a water environment to an organic environment. The NPS is simply a summation of the number of non-polar residues of the protein. The p function is the ratio of the volume of the polar exterior shell of the protein to the volume of its non-polar interior.

The value of these parameters have been calculated for 13 proteins from a range of thermophilic and mesophilic sources (Singleton and Amelunxen (1973)). Goldsach (1970) has calculated ϕH_{ave} for a very large number of proteins from a variety of sources. In both cases, although minor differences occurred, there was no definite correlation between any of these hydrophobicity parameters and thermal stability. The calculation of these parameters all make the same assumption, that on average the protein will maintain all polar residues on the exterior and all non-polar residues on the interior of the molecule. In addition, these functions cannot account for clusters of hydrophobic residues at or near the active site. Such hydrophobic clusters may protect the active site by absorbing thermal

energy, which subsequently 'melt' but without causing any marked conformational changes in the active site region. Stabilization of hydrophobic clusters might be further enhanced by hydrogen bonds or salt bridges being buried within them and hence insulated from the weakening effects of the polar aqueous environment. Margoliash and Schejter (1966) show that for cytochrome C, the hydrophobic cluster must remain constant over long evolutionary periods. The loss of buried stabilizing molecules within the clusters could give rise to thermally less stable molecules which were essentially identical in tertiary and quaternary structure and so whose enzymic activity is changed very little especially at lower temperatures.

If the melting of hydrophobic clusters is involved in stabilizing thermophilic proteins, one would expect that a marked conformational change would occur at a temperature of about 55 to 60°C. One would then expect to observe a sharp increase in the rate of thermal denaturation. In fact, many proteins do exhibit such behaviour and often have a broken Arrhenius plot with a discontinuity at about 55-60°C. Examples of these molecules include glyceraldehyde-3-phosphate dehydrogenase (Amelunxen et al (1970)), adenosine triphosphatase (Hachimori et al (1970)), glucose-6-phosphate isomerase (Muramatsu and Nosoh (1971)) and fructose 1,6-diphosphate aldolase (Sugimoto and Nosoh (1971)). 6PGDH from Bacillus stearothermophilus also exhibits a broken Arrhenius plot with a discontinuity at

about 53°C (Veronese et al (1982)).

In conclusion, it appears that microorganisms synthesize thermostable proteins and in general these appear to be physiochemically similar to their mesophilic counterparts. The mechanisms of thermophily are subtle and a more complete understanding will only come through a study of the three dimensional structure of these molecules at the atomic level. Data from X-ray crystallographic studies in conjunction with primary sequence analysis may give a greater insight into the phenomenon.

1.6 Reasons for Studying the Enzyme

The structure of 6PGDH is of interest in its own right for the following reasons:-

(1) The reaction catalysed by 6PGDH is a major source of cellular NADPH and the metabolic function and control of 6PGDH is a potentially useful area of study.

(2) 6PGDH, like isocitrate dehydrogenase and malic enzyme combines an oxidation-reduction reaction with a decarboxylation. The interaction of these two catalytic functions may be of interest in the understanding of enzyme catalysis.

(3) 6PGDH is dependent upon NADP⁺ as a coenzyme and is one of only four such enzymes to have been studied at high resolution (see section 1.7). The evolutionary history of the dehydrogenases is complex, especially that of the NADPH dependent enzymes (see section 1.7).

In addition to the above points, a comparison of the structure of the enzyme extracted from the obligate thermophile Bacillus stearothermophilus with 6PGDH extracted from sheep liver, (Adams et al (1983)), is of interest because the bacterial form exhibits a high thermal stability and resistance to denaturants (see section 1.4).

1.7 Relationship Between Dehydrogenase Structures

Table 1.3 shows details of the dehydrogenase structures which have been studied to a resolution of 3 Å or better. The enzymes can be split into two main groups; firstly those for which NAD⁺ acts as a coenzyme, i.e. Lactate dehydrogenase (LDH) (White et al (1976)), s-Malate dehydrogenase (s-MDH) (Hill et al (1972)), Liver alcohol dehydrogenase (LADH) (Eklund et al (1976)) and Glyceraldehyde phosphate dehydrogenase (GAPDH) (Biesecker et al (1977)). The rest are all NADP⁺ dependent, i.e. Dihydrofolate reductase (DHFR) (Mathews et al (1977), (1978)), Glutathione reductase (GR) (Schulz et al (1978)), p-hydroxybenzoate hydroxylase (pHBH) (Wierenga et al (1979))

and 6-Phosphogluconate dehydrogenase (6PGDH) (Adams et al (1983)).

The only difference between NADH and NADPH is the replacement of the 2' hydroxyl on the ribose of the adenosine moiety by a phosphate group (see figure 1.2). Whereas the function of the NADH-linked dehydrogenases is energy production, NADPH-linked enzymes are more varied in function and largely biosynthetic. It is not clear what role the 2' phosphate plays in this distinction.

The structure of the NAD⁺ dependent enzymes have been compared and the binding of NAD⁺ studied (e.g. Rossmann et al (1975)). The enzymes all transfer hydride directly between substrate and coenzyme and the enzyme subunits can each be described as having two domains, a coenzyme binding domain and a catalytic domain. The coenzyme binding domains all show the right-handed $\beta\alpha\beta$ connections as described by Sternberg and Thornton (1976) and the same sequence of sheet strands. Some residues important in the binding of the coenzyme are invariant; two glycines, one at the carboxyl end of sheet strand A (28 in LDH) and one just after strand D (99 in LDH), make close contacts with the coenzyme and are conserved, and an aspartate at the C-terminus of strand B (53 in LDH) which hydrogen bonds with the adenine ribose 2'-hydroxyl group is invariant. (Adams et al (1981)).

The similarity of the coenzyme binding regions of the

NAD⁺ dependent dehydrogenases has been used as evidence that these enzymes have emerged from a common precursor nucleoside binding protein. It has been suggested that this itself may have been the result of gene duplication of a mono-nucleotide binding protein. Figure 1.6 shows the fold of the sub-units for the NAD⁺ dependent dehydrogenases.

The NADP⁺ dependent dehydrogenases do not exhibit the same type of homology (the binding of coenzyme in DHFR, however, is similar, although not the same as in the NAD⁺ dependent enzymes). The fold and sub-unit structure of the NADP⁺ dependent enzymes is shown in figure 1.7. NADP⁺ dependent dehydrogenases do not even appear to exhibit convergent evolution in their binding site conformations (Adams et al 1981).

The 2'-phosphate group in DHFR interacts with an arginine residue (43) at the carboxyl end of strand B, a comparable position to the invariant aspartate of the NAD⁺ dependent dehydrogenases. The whole enzyme is however translated by one strand so that a glycine (96) at the end of strand E, which is comparable to glycine (99) of LDH, is close to the pyrophosphate and not the nicotinamide ribose.

Glutathione reductase (GR) requires both NADP⁺ and flavin adenine dinucleotide (FAD) to reduce glutathione. The conformation of NADP is similar to that of NAD⁺ in the NAD-dependent dehydrogenases (Zappe et al (1977)) but the

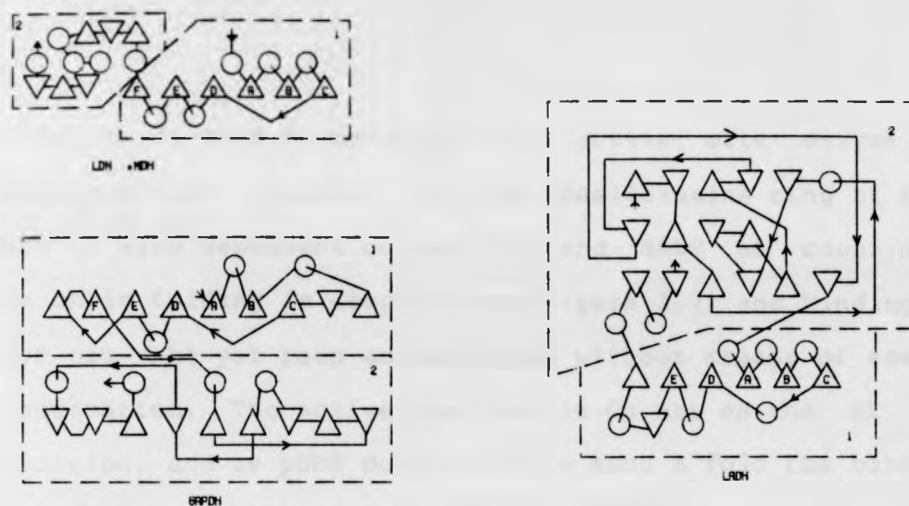


FIGURE 1.6 PROTEIN FOLDS OF THE NAD-DEPENDENT DEHYDROGENASES
CONVENTIONS ARE AS USED BY STERNBERG AND THORNTON (1977)

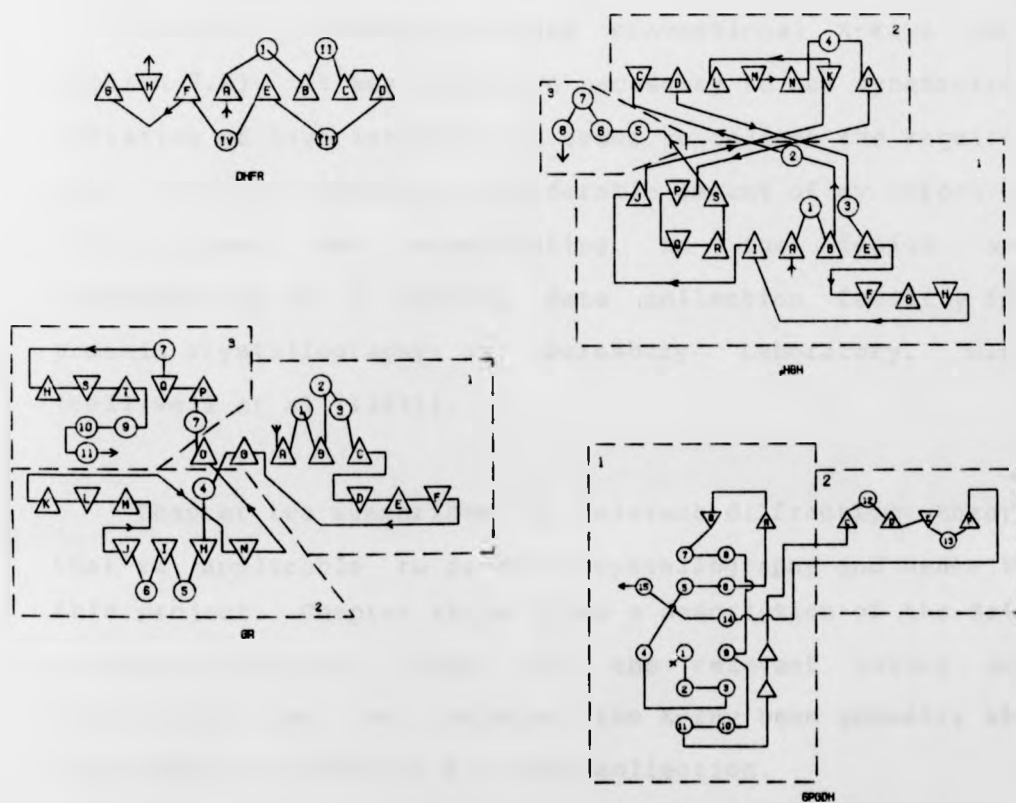


FIGURE 1.7 PROTEIN FOLDS OF THE NADP-DEPENDENT DEHYDROGENASES
CONVENTIONS AS ABOVE. 1.-COENZYME BINDING DOMAIN, 2.-CATALYTIC DOMAIN

nicotinamide must be moved in the active site enzyme to allow electron transfer to the isoalloxazine ring of FAD. pHBH is also dependent on both FAD and NADP as coenzymes. The chain folding is complex (see figure 1.7) and binding of NADP has not yet been accomplished without change of enzyme conformation. The active cystines in GR are on the II III excursion, and as pHBH does not have such a fold the binding of NADP cannot be the same in both enzymes.

1.8 Thesis Rationale

The bacterial 6PGDH crystal data could not be satisfactorily collected using conventional X-rays (see section 7.2). It was therefore necessary to use synchrotron radiation of high intensity in order to collect the required data. For this reason a considerable amount of my effort on this project was contributing to the design and commissioning of a central data collection facility for protein crystallography at Daresbury Laboratory, U.K. (Helliwell et al (1982)).

Chapter two summarises the relevant diffraction theory that is applicable to protein crystallography and hence to this project. Chapter three gives a description of the data collection facility along with the relevant theory and calculations used to optimise the X-ray beam geometry and experimental conditions for data collection.

Chapter four discusses the data collection and reduction from the bacterial 6PGDH crystals. It also contains an account of the effect of radiation damage on these crystals at different beam intensity levels.

The rationale for solving for phases was to employ molecular replacement techniques using the atomic coordinates of sheep liver 6PGDH as a model. Chapter five discusses both the theory and results of the rotation function calculations whilst chapter six details the theory and results for the translation function. The work is concluded in chapter seven with a discussion and suggestions for further work.

CHAPTER 2

CRYSTALLOGRAPHIC THEORY

2.1 Diffraction Theory

The diffraction theory applicable to protein crystallography is well documented by many authors (e.g. Blundell and Johnson (1976), Woolfson (1970)).

When a beam of X-rays interact with matter some of the electrons forming that matter act as secondary emitters of radiation. This secondary emission of radiation is known as scattering. There are two types of scattering, the first type of scattering is due to the particle nature of photons and is essentially a billiard ball effect. The incident photon collides with a comparatively loosely bound electron and is deflected from its path with some loss of energy. This loss of energy results in an increase in wavelength of the emitted photon. This incoherent type of scattering is known as Compton scattering (Compton (1923)). The second type of scattering is caused by the oscillating electromagnetic field of the photon. The alternating electric vector imparts to the electron an alternating acceleration. Classical theory (Larmor (1897)) tells us that accelerating charged particles emit electromagnetic

radiation. This secondary scattered ray is of the same wavelength as the incident wave but differs in phase by 180° . All the scattered rays from a single electron have the same phase relationship to the incident beam. This type of scattering is therefore known as coherent or Thomson scattering.

J.J. Thomson developed the theory for this type of scattering and showed that the scattering intensity $I_{2\theta}$, at angle 2θ from a beam of unpolarised X-rays is given by:

$$I = I_0 \frac{ne^4}{2r^2 m^2 c^4} (1 + \cos^2(2\theta))$$

where I_0 = incident beam intensity,

n = effective number of independently scattering electrons,

r = distance from the scatterer,

e = charge of an electron,

m = mass of an electron,

c = speed of light.

The $(1 + \cos^2 2\theta)$ term represents the partial polarisation of scattered X-rays. In X-ray crystallography Thomson scattering is enhanced by cooperative scattering from many atoms and is far more significant than the sum of the incoherent contributions.

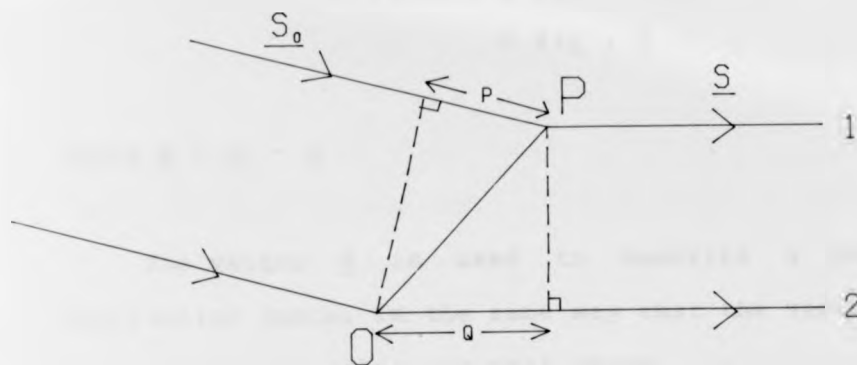


FIGURE 2.1 SCATTERING AT A POINT P RELATIVE TO AN ORIGIN O

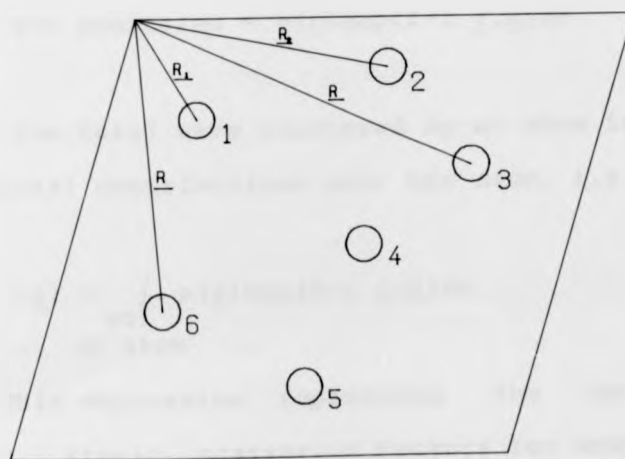


FIGURE 2.2 THE POSITION OF ATOMS IN A UNIT CELL

2.2 The Atomic Scattering Factor

The formula given above for the scattered intensity due to Thomson scattering was derived from theory based on a free point electron. The electrons in an atom, however, occupy a finite volume and are bound in certain well defined energy states. In order to get an expression for the scattering of X-rays by an atom we need to take account of their spatial distribution.

Let the electron density at a distance \underline{r} from the centre of the atom be $\rho(\underline{r})$. Consider the wave scattered at a position \underline{r} in a direction \underline{s} , relative to the wave scattered by an electron at the centre of the atom. The total wave scattered depends on the phase difference between the scattered waves.

Let us define the direction of incident radiation by a vector \underline{s}_0 and the direction of the scattered radiation by a vector \underline{s} . For simplicity in later equations let $|\underline{s}_0| = |\underline{s}| = 1/\lambda$, λ is the wavelength of the incident radiation.

From figure 2.1 we see that the path difference between ray 1 and ray 2 is $p-q$. As

$$p = \lambda \underline{r} \cdot \underline{s}_0 \quad \text{and} \quad q = \lambda \underline{r} \cdot \underline{s}$$

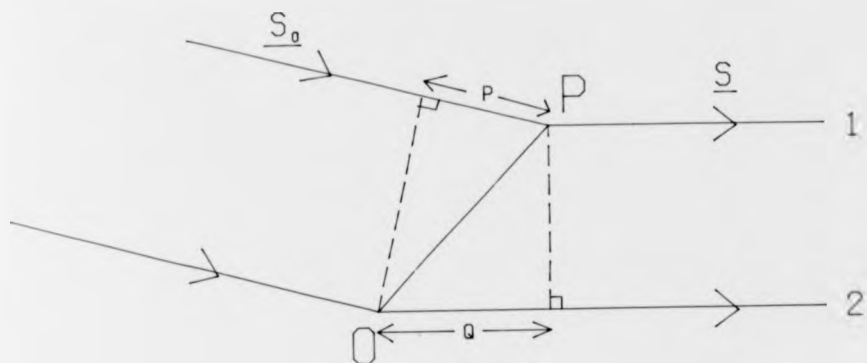


FIGURE 2.1 SCATTERING AT A POINT P RELATIVE TO AN ORIGIN O

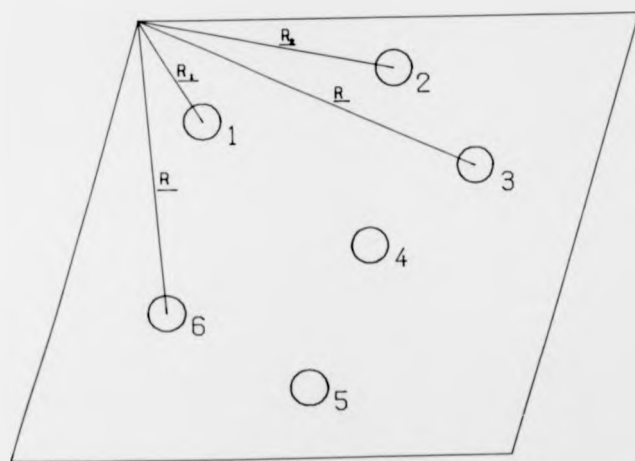


FIGURE 2.2 THE POSITION OF ATOMS IN A UNIT CELL

also the phase difference = $2\pi/\lambda \times \text{path difference}$

$$\begin{aligned}\text{therefore phase difference} &= 2\pi \underline{r} \cdot (\underline{s}_0 - \underline{s}) \\ &= 2\pi \underline{r} \cdot \underline{S}\end{aligned}$$

where $\underline{S} = \underline{s}_0 - \underline{s}$

The vector \underline{S} is used to describe a position in diffraction space, in the same way that the vector \underline{r} is used to describe a position in real space.

The total wave scattered by a small volume element dv at a position \underline{r} relative to the wave scattered at the origin will therefore have an amplitude proportional to $\rho(\underline{r})dv$ and phase $2\pi \underline{r} \cdot \underline{S}$, i.e.

$$\text{wave scattered} = \rho(\underline{r}) \exp(2\pi i \underline{r} \cdot \underline{S}) dv$$

Hence the total wave scattered by an atom is the integral of individual contributions over the atom, i.e.

$$F(\underline{S}) = \int_{\substack{\text{vol} \\ \text{of atom}}} \rho(\underline{r}) \exp(2\pi i \underline{r} \cdot \underline{S}) dv$$

This expression represents the atomic scattering factor. Atomic scattering factors for most atoms have been calculated and are tabulated in International Tables for X-ray Crystallography Volume III (1968). The above expression for the atomic scattering factor has been derived

for an atom at rest. Debye (1914) showed that due to thermal vibrations the scattered wave is decreased in intensity by a factor of:

$$\exp[-B \sin^2 \theta / \lambda^2]$$

where $B = 8\pi^2 \bar{u}^2 / 3$

and \bar{u} = mean displacement of atom along normal to the reflecting planes.

The above relationships have been based on the normal assumptions that the electron density and thermal vibrations are isotropic. Temperature factors are discussed in Chapter 4.

2.3 The Molecular Transform

We now derive the expression for the scattering by an assembly of atoms placed at defined positions in a unit cell.

Let us consider atom 1 in figure 2.2 which is at a distance \underline{r}_1 from the origin. This shift in origin from the centre of the atom means that the distance \underline{r} in the equation for the scattering of an atom becomes $\underline{r} + \underline{r}_1$. Hence the scattering by an atom 1 relative to the new origin is:

$$\mathcal{J} = \int_{\substack{\text{vol} \\ \text{of atom}}} \rho(\underline{r}) \exp(2\pi i(\underline{r}_1 + \underline{r}) \cdot \underline{S}) d\underline{v}$$

$$\text{Therefore } \mathcal{J}_1 = f_1 \exp(2\pi i \underline{r}_1 \cdot \underline{S})$$

$$\text{where } f_1 = \int_{\substack{\text{vol} \\ \text{of atom}}} \rho(\underline{r}) \exp(2\pi i \underline{r} \cdot \underline{S}) d\underline{v}$$

Similar expressions may be deduced for atoms 2,3 and all the other atoms in the unit cell.

The total wave scattered by all the atoms is given by the vector sum of the individual contributions from each of the atoms. Hence the total wave scattered is given by:

$$\underline{G}(\underline{S}) = \mathcal{J}_1 + \mathcal{J}_2 + \mathcal{J}_3 + \dots + \mathcal{J}_N$$

i.e.

$$\underline{G}(\underline{S}) = \sum_{j=1}^N f_j \exp(2\pi i \underline{r}_j \cdot \underline{S})$$

This equation represents the molecular transform. The function is complex and varies continuously over all diffraction space \underline{S} .

2.4 Structure Factor Equation

If we consider a one dimensional crystal with a linear array of unit cells of repeat distance a , the total wave scattered will be the sum of the waves scattered by each unit cell. As we have just shown, the wave scattered due to the first unit cell relative to the origin is simply $\underline{G}(\underline{S})$.

The wave scattered by the second unit cell relative to the same origin is $\underline{G}(\underline{S})\exp(2\pi i \underline{a} \cdot \underline{S})$ since all the distances are shifted by the vector \underline{a} . The scattered wave for the n unit cell is therefore $\underline{G}(\underline{S})\exp(2\pi i (n-1)\underline{a} \cdot \underline{S})$. Hence the total wave scattered is

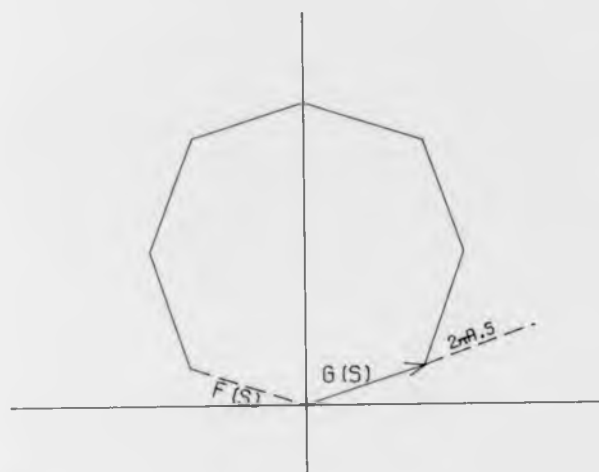
$$\underline{F}(\underline{S}) = \sum_{n=1}^T \underline{G}(\underline{S})\exp(2\pi i (n-1)\underline{a} \cdot \underline{S})$$

where T = total number of unit cells.

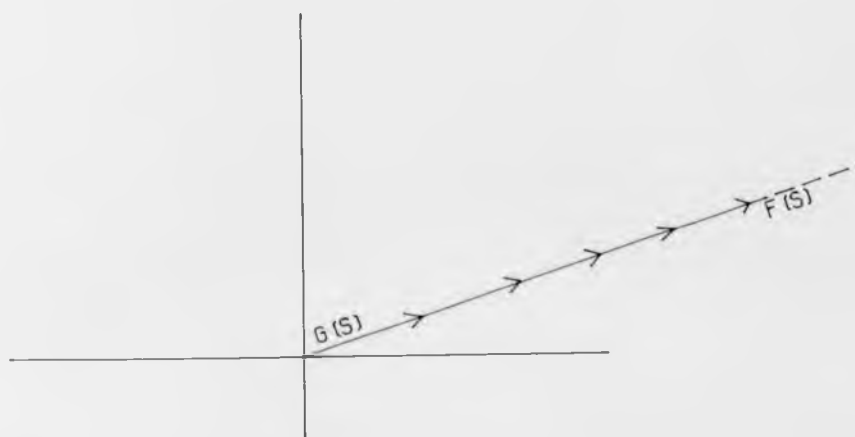
The way in which these individual contributions add up may be seen from the vector diagram (figure 2.3). The wave from each unit cell is out of phase with its neighbour by an amount $2\pi \underline{a} \cdot \underline{S}$. Hence as the number of cells become large, the total wave scattered, $\underline{F}(\underline{S})$, will be approximately the same order of magnitude as $\underline{G}(\underline{S})$, the molecular transform. The molecular transform for X-rays is too small to be observed (Blundell & Johnson (1976)). How then is scattering observed from a crystal?

Scattering will only be observed when the phase difference between the waves scattered by successive unit cells is equal to an integral number of 2π (see figure 2.3(b)). i.e.

$$\begin{aligned} 2\pi \underline{a} \cdot \underline{S} &= 2\pi h \\ \Rightarrow \underline{a} \cdot \underline{S} &= h \quad \text{where } h \text{ is an integer} \end{aligned}$$



(A)



(B)

FIGURE 2.3 VECTOR DIAGRAMS TO ILLUSTRATE TOTAL
 WAVE SCATTERED BY A CRYSTAL. (A) THE PHASE DIFFERENCE
 BETWEEN WAVES FROM ADJACENT UNIT CELLS IS $2\pi A.S.$.
 (B) THE PHASE DIFFERENCE IS AN INTEGRAL MULTIPLE
 OF 2π

Under these circumstances the waves add up constructively to form a significant scattered wave which is proportional in magnitude to $T \times |G(\underline{S})|$. For a crystal 0.15 mm long with a unit cell of 120\AA (similar to bacterial 6PGDH) $T = 1.25 \times 10^4$ and hence the amplification is significant.

In summary, for a one dimensional lattice, scattering will only be observed in a particular direction \underline{S} where \underline{S} satisfies the equation $\underline{a} \cdot \underline{S} = h$. When the problem is extended to three dimensions with a unit cell defined by vectors $\underline{a}, \underline{b}$ and \underline{c} the condition for diffraction becomes:

$$\underline{a} \cdot \underline{S} = h$$

$$\underline{b} \cdot \underline{S} = k$$

$$\underline{c} \cdot \underline{S} = l$$

where h, k and l are integers. These equations are known as the Laue equations (Laue (1912, 1913)). They enable us to rewrite the total wave scattered by a crystal in a direction \underline{S} in the following way:

$$\underline{F}(\underline{S}) = \sum_{j=1}^N f_j \exp(2\pi i \underline{r}_j \cdot \underline{S})$$

neglecting the proportionality constant T . Let the fractional coordinates of the j^{th} atom be x_j, y_j, z_j i.e.

$$\underline{r}_j = \underline{a}x_j + \underline{b}y_j + \underline{c}z_j$$

Hence

$$\begin{aligned}\underline{r}_j \cdot \underline{S} &= x_j \underline{a} \cdot \underline{S} + y_j \underline{b} \cdot \underline{S} + z_j \underline{c} \cdot \underline{S} \\ &= hx_j + ky_j + lz_j \quad (\text{from Laue's equations}).\end{aligned}$$

Therefore

$$\underline{F}(\underline{hkl}) = \sum_{j=1}^N f_j \exp(2\pi i (hx_j + ky_j + lz_j))$$

where \underline{S} on the left hand side has been replaced by \underline{hkl} .

This equation is known as the Structure Factor equation. It represents the molecular transform sampled at the reciprocal lattice points \underline{hkl} (see section 2.1.4). If the positions of all atoms are known then the corresponding diffraction pattern can be calculated.

The structure factor is a complex quantity and is often expressed in terms of its amplitude and phase. An equivalent expression for structure factor equations is

$$\underline{F}(\underline{hkl}) = F(\underline{hkl}) \exp i\alpha(\underline{hkl})$$

This can be expanded into its real and imaginary parts, i.e.

$$\underline{F}(\underline{hkl}) = A + iB$$

$$\text{where} \quad A = F(\underline{hkl}) \cos \alpha(\underline{hkl})$$

$$\text{and} \quad B = F(\underline{hkl}) \sin \alpha(\underline{hkl})$$

This is represented on an Argand diagram (see figure 2.4).

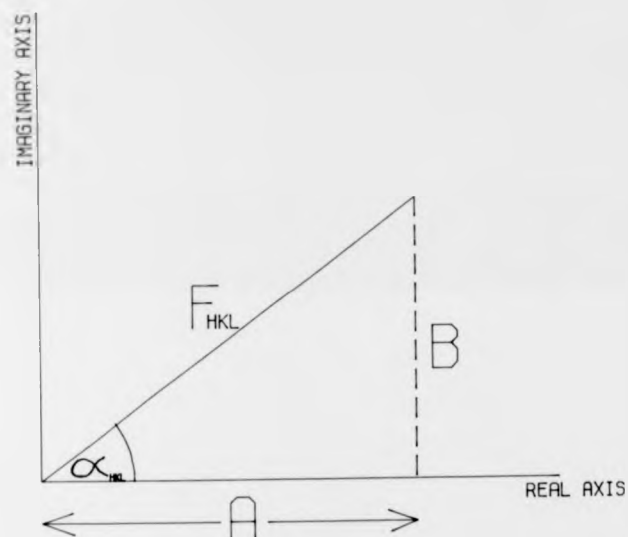


FIGURE 2.4 ARGAND DIAGRAM SHOWING THE REAL
AND IMAGINARY PARTS OF THE STRUCTURE FACTOR F_{HKL}

When the scattered X-radiation is recorded on photographic film or scintillation counter, all information on the phase is lost and only a measurement of energy, that is the intensity of the diffracted ray, is recorded. The intensity is given by:

$$\begin{aligned} I(hkl) &= \underline{F}(hkl) \cdot \underline{F}^*(hkl) \\ &= F(hkl)^2 \end{aligned}$$

where \underline{F}^* = the complex conjugate of \underline{F}

2.5 Reciprocal Lattice and Ewald's Construction

To understand fully the implications of the Structure Factor equation we must make use of the Convolution theory of Fourier analysis. The theorem states:

"The Fourier transform of the convolution of two functions is the product of their Fourier Transforms."

The converse theorem states:

"The Fourier transform of the product of two functions is the convolution of the transform of one with the transform of the other."

A crystal is the result of the convolution of a molecule with a lattice (i.e. if we take the molecule and place it at each of the lattice points in turn we will build up to a crystal). Hence the diffraction pattern of the molecular crystal is the product of the diffraction pattern

of the molecule (molecular transform) with the Fourier transform (diffraction pattern) of the lattice.

The diffraction pattern of a lattice is also a lattice but one whose dimensions are inversely proportional to the dimensions of the real lattice (the reciprocal lattice, see figure 2.5). The product of these two transforms results in a sampling of the molecular transform at each of the reciprocal lattice points.

Hence we see that the lattice nature of a diffraction pattern arises from the crystal lattice and that the intensity of each spot is governed by the intensity at that point of the underlying molecular transform. In general, each part of the molecule contributes to every part of the diffraction pattern. Conversely in order to reconstruct the molecule from its diffraction pattern, it is necessary to measure the intensities of all the diffraction spots.

The diffraction vector \underline{S} that we have been using may be specified in terms of the reciprocal cell constants \underline{a}^* , \underline{b}^* and \underline{c}^* . i.e.

$$\underline{S} = h\underline{a}^* + k\underline{b}^* + l\underline{c}^*$$

Diffraction will only occur when h , k and l are integers. As Bragg's Law tells us that at this time the diffracted beam must be scattered through an angle 2θ , where θ satisfies the equation

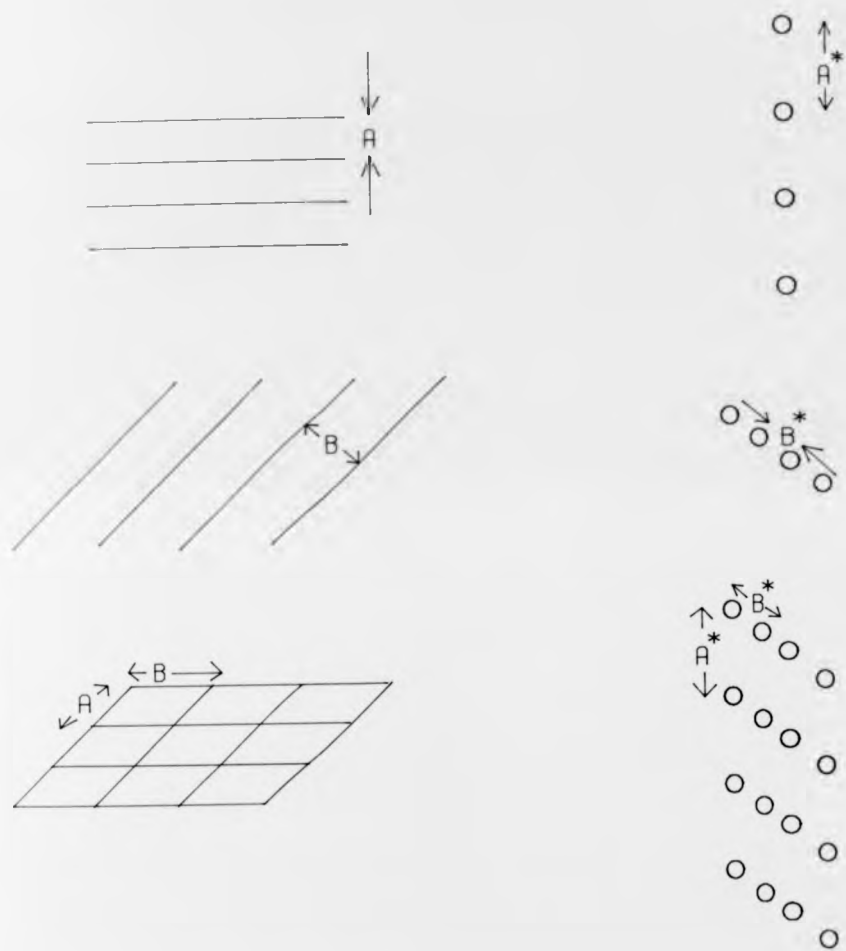


FIGURE 2.5 THE TRANSFORMS OF A SET OF LINES AND OF A LATTICE

$$2d \sin \theta = n \lambda$$

We can therefore see that

$$\underline{S} = 2 \sin \theta / \lambda = 1/d \text{ for } n = 1$$

i.e. The magnitude of the vector S which defines a position in diffraction space is inversely proportional to the interplanar spacing of the planes that give rise to that particular diffraction peak.

From Laue's equations

$$\underline{a} \cdot \underline{S} = h$$

$$\underline{b} \cdot \underline{S} = k$$

$$\underline{c} \cdot \underline{S} = l$$

Hence:

$$\underline{a} \cdot \underline{a}^* = 1 \quad \underline{a} \cdot \underline{b}^* = 0 \quad \underline{a} \cdot \underline{c}^* = 0$$

$$\underline{b} \cdot \underline{a}^* = 0 \quad \underline{b} \cdot \underline{b}^* = 1 \quad \underline{b} \cdot \underline{c}^* = 0$$

$$\underline{c} \cdot \underline{a}^* = 0 \quad \underline{c} \cdot \underline{b}^* = 0 \quad \underline{c} \cdot \underline{c}^* = 1$$

From these equations it is apparent that \underline{a}^* and \underline{b}^* are perpendicular to the real c axis. Conversely, \underline{c}^* is perpendicular to a and b. The relationships between the real and reciprocal unit cell constants for the various crystal systems are given in International Tables for X-ray Crystallography Vol. I (1968).

The condition for diffraction of X-rays by a crystal may be expressed either in terms of Bragg's Law, or in terms of the Laue equations. Ewald (1921) proposed a simple geometrical construction which encompasses both of these laws.

A sphere is drawn with centre at the crystal (c) and radius $1/\lambda$. The origin of the reciprocal lattice is placed at a point O where the unreflected X-ray beam travelling in a direction AC meets the sphere (see figure 2.6). The condition that a particular ray CB, is a diffracted ray may then be expressed as: X-rays will be diffracted in the direction CB if the point B represents a reciprocal lattice point (hkl) where hkl are integers, i.e. the vector OB is a reciprocal lattice vector $\underline{S} = h\underline{a}^* + k\underline{b}^* + l\underline{c}^*$.

Hence, in order to bring the reflection hkl into the diffracting position the crystal must be rotated in such a way that this particular lattice point cuts the sphere of reflection. Reciprocal lattice vectors have dimensions of \AA^{-1} , however they can be made dimensionless by allowing the sphere of reflection to equal unity and a scale factor equal to the incident wavelength taken into the calculation of the reciprocal lattice dimensions.

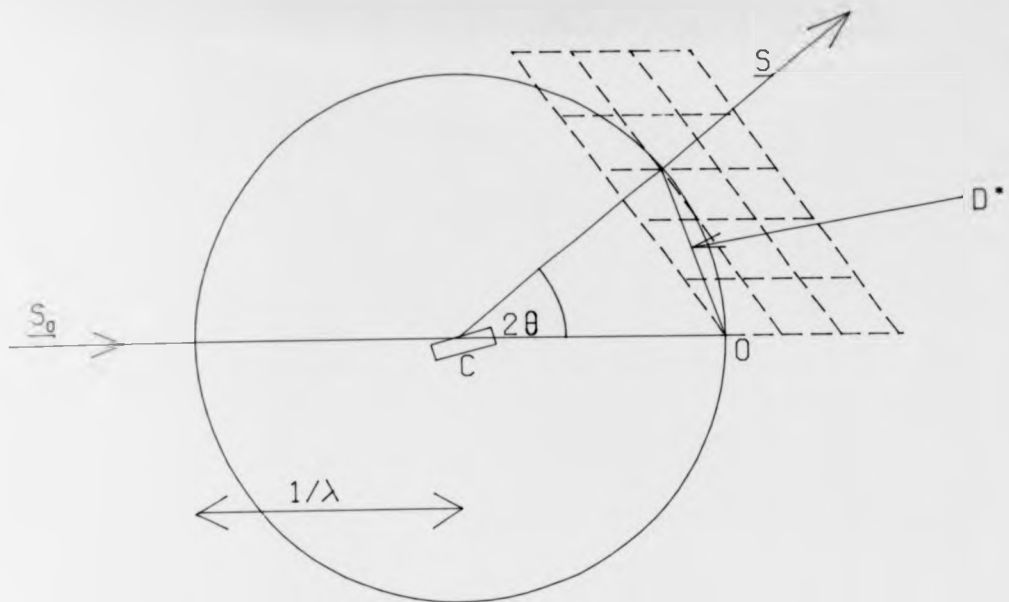


FIGURE 2.6 EWALD'S CONSTRUCTION

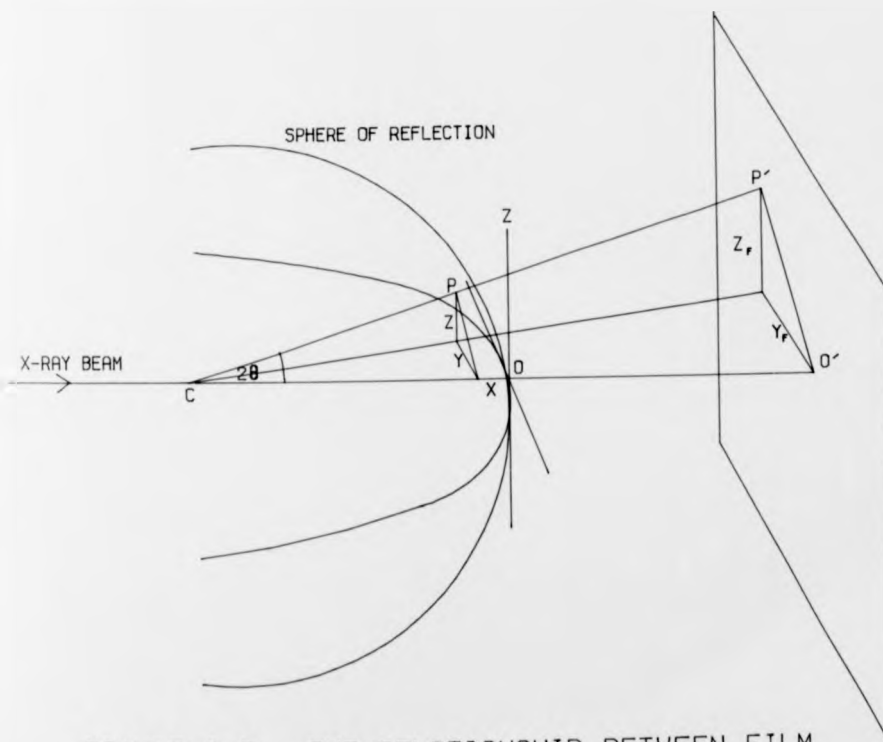


FIGURE 2.7 THE RELATIONSHIP BETWEEN FILM COORDINATES (X_F, Y_F) AND CRYSTAL COORDINATES (X, Y, Z)

2.6 The Electron Density Equation

We have seen in section 2.1.3 that we can predict the diffraction pattern from the structure. How do we then predict the structure from the diffraction pattern? Again the answer lies in Fourier Transform theory. The diffraction pattern is the Fourier transform of the structure. It therefore follows that the structure is the Fourier transform of the diffraction pattern.

In order to show this, the structure factor equation is rewritten in terms of a continuous summation over the volume of the unit cell, i.e.

$$\begin{aligned} \underline{F}(\underline{S}) &= \sum_{j=1}^N f_j \exp(2\pi i \underline{r}_j \cdot \underline{S}) \\ &= \int_{\text{volume of unit cell}} \rho(\underline{r}) \exp(2\pi i \underline{r} \cdot \underline{S}) d\underline{v} \end{aligned}$$

By multiplying both sides by $\exp -2\pi i \underline{r} \cdot \underline{S}$ and integrating over the volume it may be shown that

$$\rho(\underline{r}) = \int_{\text{volume of unit cell}} \underline{F}(\underline{s}) \exp(-2\pi i \underline{r} \cdot \underline{S}) d\underline{v}_s$$

where $d\underline{v}_s$ is a small unit of diffraction space.

The integration can be replaced by a summation since $\underline{F}(\underline{s})$ is not continuous and is non zero only at the reciprocal lattice points. Hence:

$$\rho(xyz) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \underline{F}(hkl) \exp(-2\pi i(hx + ky + lz))$$

Thus if the structure factors $\underline{F}(hkl)$ are known for all reflections then the electron density may be calculated for for each point xyz in the unit cell. By identifying particular atoms and residues from their electron density the three dimensional structure can be solved.

2.7 The Phase Problem

In order to calculate the electron density we need to know $\underline{F}(s)$, i.e. both the amplitude and phase of the structure factors. This can be emphasized by rewriting the electron density equation as follows:

$$\rho(xyz) = \frac{1}{V} \sum_{h=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} F(hkl) \exp i\alpha(hkl) \exp -2\pi i(hk+ky+lz)$$

In the recorded intensities obtained on film or scintillation counter, only the intensities, and hence amplitudes, are measured; all information on phase is lost. It can therefore be seen that direct structure determination from recorded intensities is impossible since part of the information is missing. The problem of phase determination is the basic problem in any crystal structure determination.

There are five methods by which the phase problem may

be overcome, i.e.

- (1) The Patterson summation,
- (2) Direct methods,
- (3) Heavy atom isomorphous replacement,
- (4) Anomalous scattering,
- (5) Molecular replacement.

(1) The Patterson summation is a Fourier summation based on the square of observed structure factor amplitudes and is essentially a vector map of the structure. For molecules containing relatively few atoms (<25) this can often be interpreted in terms of an approximation to the structure. In larger molecules containing heavy atoms the position of the heavy atom sites may be elucidated by means of their Patterson maps.

(2) Direct methods involve using statistical and mathematical probability relationships between reflections to obtain phase information. This again is limited to smaller molecules.

(3) The method of heavy atom isomorphous replacement in which a heavy atom (and hence strong scattering) is introduced into a light atom structure is the most commonly used method in protein crystallography for obtaining phases.

(4) Anomalous scattering in which phase information is obtained from intensity changes in the diffraction pattern caused by absorption of radiation by certain atoms whose natural frequency of vibration is close to that of the

incident radiation.

(5) Molecular replacement techniques consist of trying to superimpose 'known structures' of related molecules onto the structure under investigation. If a large overlap occurs between the known and unknown structures in Patterson space, the initial phases for the unknown structure may be calculated using the atomic coordinates of the known structure. It is this last method by which we will try to phase the bacterial 6PGDH data.

2.8 The Rotation Method

The rotation method has been very well documented in a book edited by Arndt and Wonacott (1977). The camera used is of a simple design with a collimated X-ray beam hitting a crystal which is mounted on a spindle axis perpendicular to the X-ray beam. The crystal is usually mounted so that one of its real lattice axes is coaxial with the camera rotation axis. The reflections are recorded on a flat film as the crystal is rotated through a given angle. The geometry of the camera in relation to the film and reciprocal lattice is shown in figure 2.7. The reciprocal lattice coordinates (xyz) at the point of intersection (P) of the Ewald sphere are related to the indices of the reflection (hkl) by an expression of the form:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [\phi] [\psi_z] [\psi_y] [\psi_x] [A]$$

where $[\psi_x]$, $[\psi_y]$ and $[\psi_z]$ are (3x3) rotational matrices defining the small misorientations about the 'perfectly set' crystal axes.

$[\phi]$ is a rotation matrix corresponding to the spindle axes.

$[A]$ is a rotation matrix defining the standard orientation of the reciprocal lattice vectors before rotation.

CHAPTER 3

EXPERIMENTAL SET-UP AND DATA COLLECTION

3.1 Introduction

The data used for the work contained within this thesis was collected on the Protein Crystallography workstation on beamline seven at the S.E.R.C. Synchrotron Radiation Source (SRS) at Daresbury, U.K.

Four native crystals were used to collect sixty degrees (four equivalents) of data using a modified Ardent-Wonacott Rotation camera and CEA film. The enzyme was extracted and crudely purified at the Microbiological Establishment, Porton, Wilts, U.K. (Atkinson et al (1972)). Further purification and crystallisation was done by B.M.F. Pearse and J.I. Harris (1973) at the Medical Research Council Laboratory of Molecular Biology, Cambridge.

Two earlier datasets - a native and a platinum derivative - were collected on the double focussing X-ray camera used for protein crystallography at DORIS/EMBL in Hamburg which is described by Bartunik et al (1982). These unfortunately proved difficult to process for two main reasons. Firstly the crystal to film distance employed when

the data was being collected was too small, causing indexing problems of the diffraction spots. This distance was not increased due to the instability of the camera alignment system. Secondly, and probably more importantly, the still photographs which are used to calculate the crystal misorientation matrices were taken at the current spindle setting after every second data photograph rather than at 0° and 90° before and after data collection. It was thought that any crystal slippage would be more easily pinpointed by this method. Unfortunately this led to a rather limited angular separation between the start and end stills causing problems with the misorientation matrix determination algorithms for this large unit cell system.

3.2 Synchrotron Radiation

Lienard (1898) demonstrated that an electron which follows a circular orbit at constant speed is constantly changing its direction and therefore accelerated by a centripetal force. As accelerated charged particles emit electromagnetic radiation (Larmor (1897)) an electron moving in a circular orbit is an intense source of e.m. radiation.

In synchrotrons and synchrotron storage rings electrons are accelerated to speeds very close to the speed of light and kept circulating around a closed orbit by means of powerful magnetic fields. One result of this is that a

broad spectrum of electromagnetic radiation (synchrotron radiation) is emitted at a tangent to the electron orbit. The physical properties of the emitted synchrotron radiation (SR) are well understood and calculable from the basic machine parameters (Schwinger 1946, 1949, 1954). The most relevant properties are reviewed by several authors (e.g. Winick and Bienenstock (1978), Materlik (1982)). The spectrum of emitted SR from the SRS is shown in figure 3.1

The main properties of SR relevant to crystallography are:

- (1) Intense photon flux;
- (2) Smoothly varying spectral profile over a large range of wavelengths;
- (3) High collimation of the X-ray beam.

Other potentially useful properties are:

- (4) Large degree of polarisation of the X-ray beam;
- (5) Pulsed time structure.

The intense photon flux obtained from a synchrotron source is several orders of magnitude larger than that obtained from a 'conventional' X-ray source. This intense flux means that data can be collected much more quickly. Typical reduction in exposure times are factors of 40-125 (Wilson et al (1983)). It has also been observed that for most proteins an increase in the effective crystal lifetime of the order of two to six times is obtained using SR (Wilson et al (1983)).

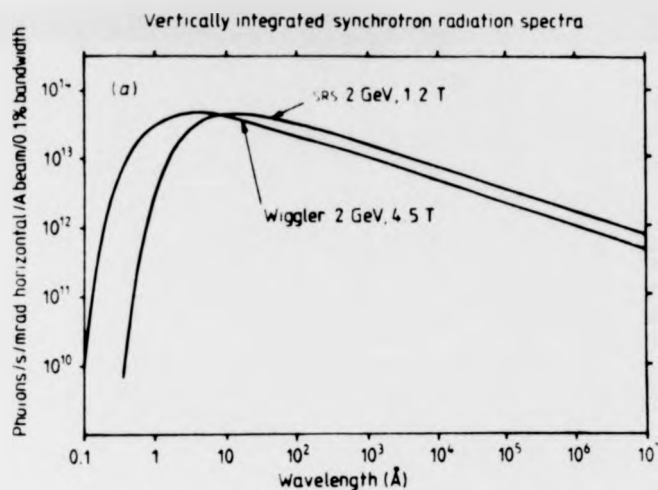


FIGURE 3.1 THE INTENSITY SPECTRUM OF THE SRS ON A CONVENTIONAL AND WIGGLER MAGNET

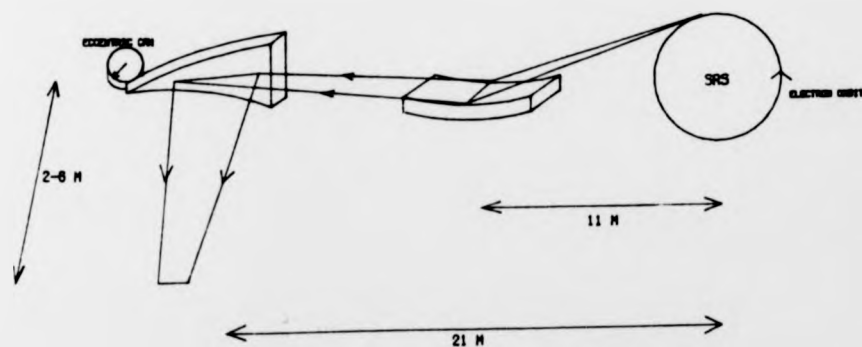


FIGURE 3.2 SCHEMATIC DIAGRAM SHOWING THE PRINCIPLE OF THE FOCUSING AND MONOCHROMATOR SYSTEM

In February 1982 a vertically focussing mirror was installed in beamline 7.2 of the SRS with a gain in intensity of three to four times. Some workers (Dodson et al (1983), Achari et al (1983), Wonacott pers. comm.) report that data collected since then shows a reduction in effective crystal lifetime when compared to synchrotron data collected prior to the mirror installation. Further discussion of radiation damage of biological samples, in particular bacterial 6PGDH both prior to and after the mirror installation is given in chapter four.

The smoothly varying spectral profile means that the wavelength of the X-rays used for protein crystallography can be chosen to be any one of a large number of absorption edges. This allows anomalous scattering to occur and subsequent solution of the phase problem. Although in theory one could do this with the Bremstrahlung radiation of a conventional source, the reduction in source brightness away from the characteristic wavelength proves too inhibitory. Synchrotron Radiation is approximately 10^5 times more powerful than conventional X-ray Bremstrahlung.

The high collimation of the X-ray beam due to the convergent beam geometry was essential for our small crystals. An average crystal did not exceed 0.15 mm in any dimension.

3.3 The mirror

All the vertical focussing of the synchrotron beam is achieved by means of a platinum coated mirror. The mirror is made of fused quartz, polished and coated with platinum under UHV conditions. The dimensions of the block are 582 mm long, 80 mm wide and 30 mm deep. The object distance, p_v is fixed at 11 m (the distance from the ring to the mirror, see figure 3.2). The image distance, p_v' can be varied from 12-17 m.

The mirror is initially flat but is bent to produce a cylindrical surface with bending radius, R , of approximately 1 km, by means of a couple applied at each end. The exact bending radius is given by the following:

$$\frac{2}{R} = \frac{\sin \theta_c}{p_v} + \frac{\sin \theta_c}{p_v'}$$

where θ_c = the critical angle of reflection.

The critical angle of reflection is the lowest incident angle at which reflection can occur and is variable with wavelength. For platinum it follows the relationship:

$$\theta_c \text{ (radians)} \approx 0.006 \lambda^{\circ} \text{ (Angstroms)}$$

The horizontal divergence is unaffected by reflection from the mirror. The vertical image size is given by:

$$h_v (p'_v/p_v)$$

where h_v = vertical source size.

For $p'_v = 12-17$ m (as at the SRS) the image size is in the range 0.47-0.67 mm (FWHM) with a convergence angle given by:

$$\beta_v (p_v/p'_v)$$

where β_v = the incident vertical divergence.

The convergence angle at 1.5 \AA for the SRS is in the range 0.23-0.165 mrad. The mirror sets a lower limit on the available range of wavelengths at the workstation of approximately 1 \AA . Absorption by the Beryllium windows sets the upper limit at about 3 \AA . The surface roughness of the block has been measured as 8.7 \AA rms over the whole surface (J. Bennett (1981) pers. comm.).

3.4 Monochromatisation of the X-ray Beam

A curved, single, bent, crystal monochromator as described by Helliwell et al (1982) was used both to monochromate and focus (in the horizontal) the synchrotron beam. The principle of the system is shown in figure 3.2

The (111) layers of germanium were cut at an angle, α , of 10.5° to the surface of the crystal producing a compression of the beam given by:

$$A = \sin (\theta - \alpha) / \sin (\theta + \alpha)$$

see figure 3.3

This compression of the beam allows more of the available synchrotron radiation to be utilised by small samples. Further 'fine tune' focussing is obtained by means of rotating the eccentric cam (see figure 3.2) which alters the curvature of the crystal surface.

3.4.1 The Guinier Position

The radius of curvature R , for a triangular plate such as our monochromator crystal is given by the formula:

$$R = Eht^3/12LW \quad (\text{Hendrix et al (1979)})$$

where E = modulus of elasticity

h = length of the base of the triangle

t = thickness

L = height of the triangle

W = force applied to tip

When one bends the monochromator crystal it obviously affects the Bragg angle at which X-rays are reflected along its length and therefore the energy (wavelength) of the

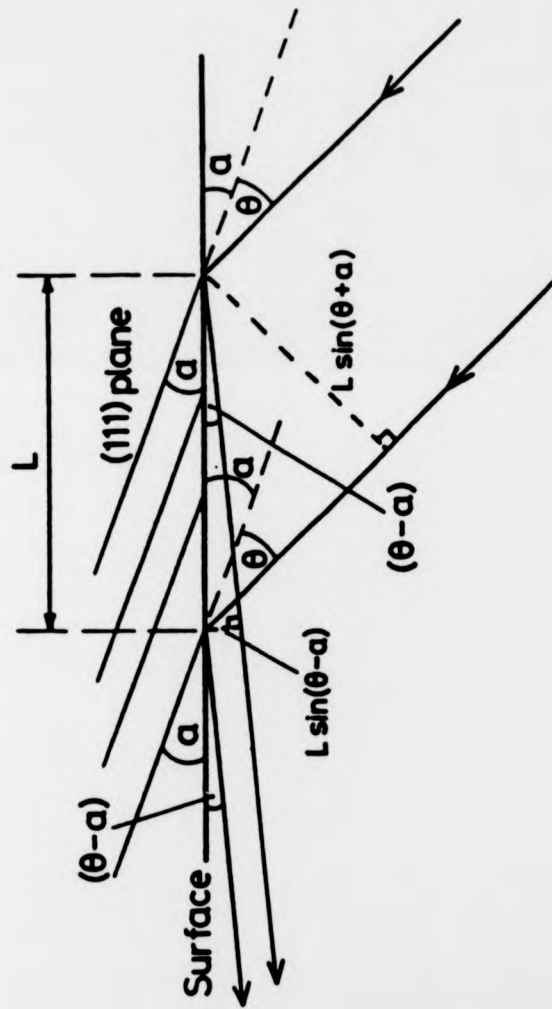
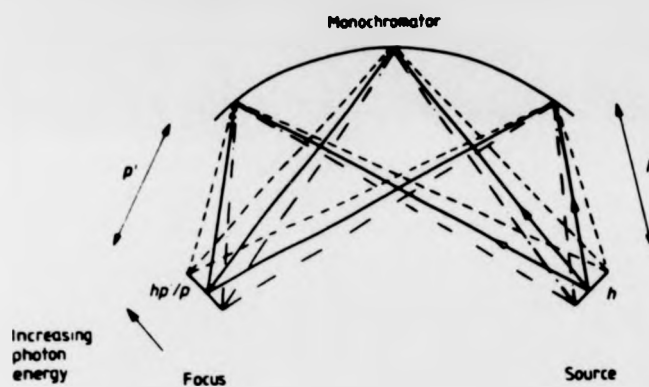
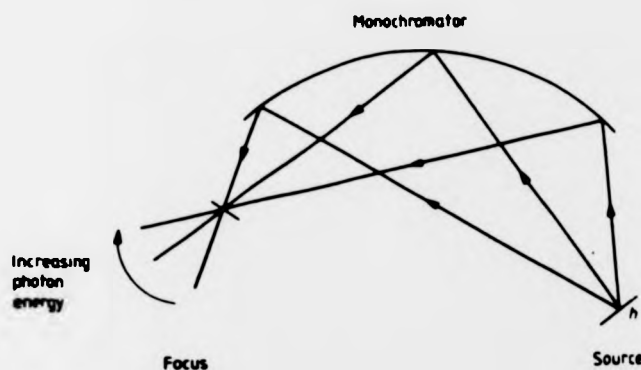


Figure 3.3 Diagram Showing Beam Compression for Reflection from an Asymmetrically Cut Crystal Monochromator



(A) For a finite horizontal source rays from any point of that source form equal angles with the monochromator surface along its length at the Guinier condition. Rays from different points of the source change the angle of incidence so that at the focus there is an energy gradient across its width. At the SRS this is ≈ 20 eV at $\lambda = 1.743$ Å (Fe K edge).



(B) When the monochromator curvature is not at the Guinier position rays from a point on the source to different points on the monochromator are incident at different angles. At a point on the focus the converging rays, from the equivalent point on the source, possess a correlation of energy with direction.

Figure 3.4 Ray Diagram Showing (A) The Guinier Condition
(B) The 'Overbent' Condition

processing. A detailed study of the effects of horizontal and vertical beam divergences and correlated $\frac{\delta\lambda}{\lambda}$'s on the recording ranges and hence spot shapes and sizes has been published by Greenhough et al (1982).

3.4.2 Spectral Resolution

The wavelength spread in the reflected beam, $\frac{\delta\lambda}{\lambda}$ is the convolution of the contribution due to the rocking width of the monochromator crystal (i.e. the range in angles that are accepted in the incident beam that are then subsequently reflected in the 'monochromatic' beam) with the contribution due to geometrical factors of source size and incident angle along the curved monochromator surface. The contribution due to the rocking width of the crystal is:

$$\left(\frac{\delta\lambda}{\lambda}\right)_{\text{crystal}} = \omega_{\text{acc}} \cot \theta$$

where ω_{acc} = acceptance angle of incident beam (rocking width)

θ = Bragg angle of reflection from monochromator.

The contribution due to the geometrical factors is discussed by Lemonnier et al (1978) and is given by:

$$\left(\frac{\delta\lambda}{\lambda}\right)_{\text{geom}} = \left[\frac{h}{b} + \frac{L}{2} \left| \frac{\sin r}{p'} - \frac{\sin i}{p} \right| \right] \cot \theta$$

where h = horizontal source size,

L = length of monochromator illuminated.

Therefore the total spectral resolution is given by:

$$\frac{\delta\lambda}{\lambda} = \cot\theta \left[\left(\frac{h}{p} + \frac{L}{2} \left| \frac{\sin r}{p'} - \frac{\sin i}{p} \right| \right)^2 + \omega_{acc}^2 \right]^{\frac{1}{2}}$$

The variation of $\frac{\delta\lambda}{\lambda}$ as a function of p' is shown in figure 3.5.

3.4.3 Horizontal Spot Width

For a conventional source the spot width is a convolution of the contributions due to four factors, viz:

- (1) The crystal size, δ_c ;
- (2) The crystal mosaicity, δ_n ;
- (3) The spectral bandwidth, $\delta_{(\delta\lambda/\lambda)}$
- (4) The electron focus of the source, δ_f

where

$$\begin{aligned}\delta_c &= \frac{x}{s} (l'_x + s) \\ \delta_f &= \frac{f}{s} l'_x \\ \delta_n &= l'_x n d^* \\ \delta_{(\delta\lambda/\lambda)} &= \frac{2l'_x \delta\theta}{x}\end{aligned}$$

and x = length of the crystal,

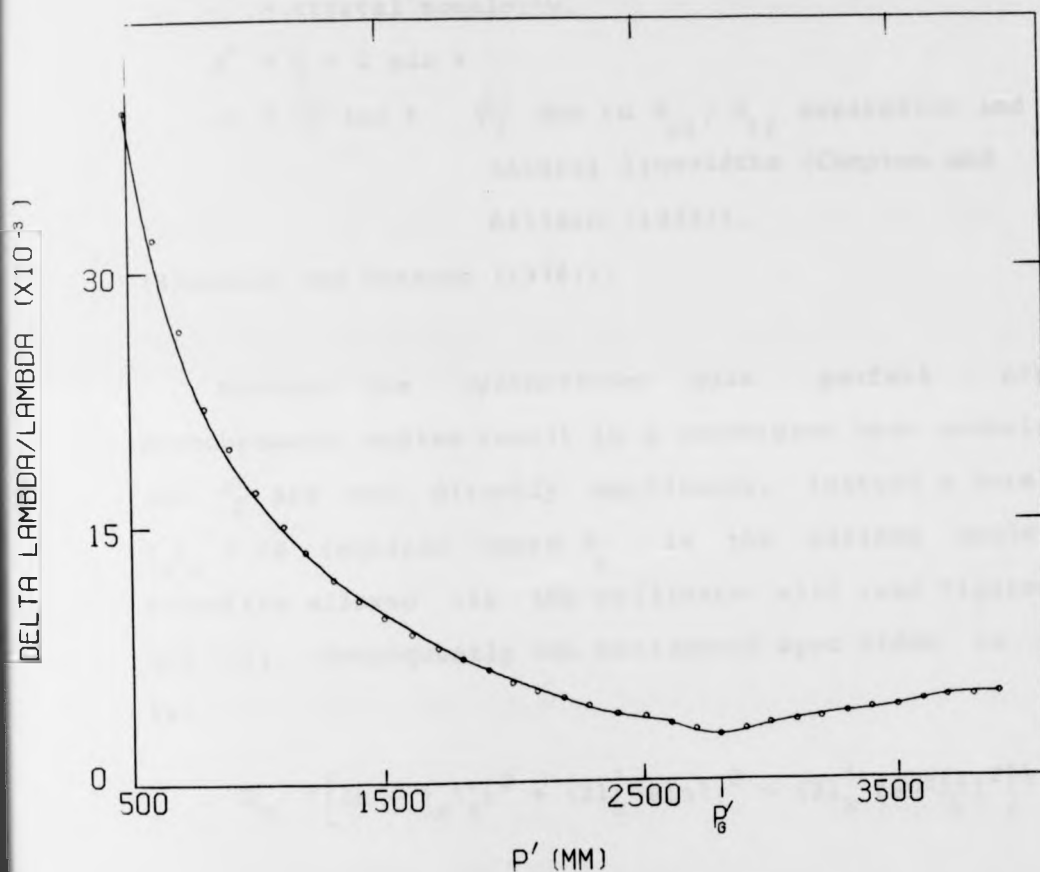


FIGURE 3.5 SPECTRAL RESOLUTION AS A FUNCTION OF P' FOR A GERMANIUM MONOCHROMATOR ON THE SRS ($\lambda=1.488$ ANGSTROMS)

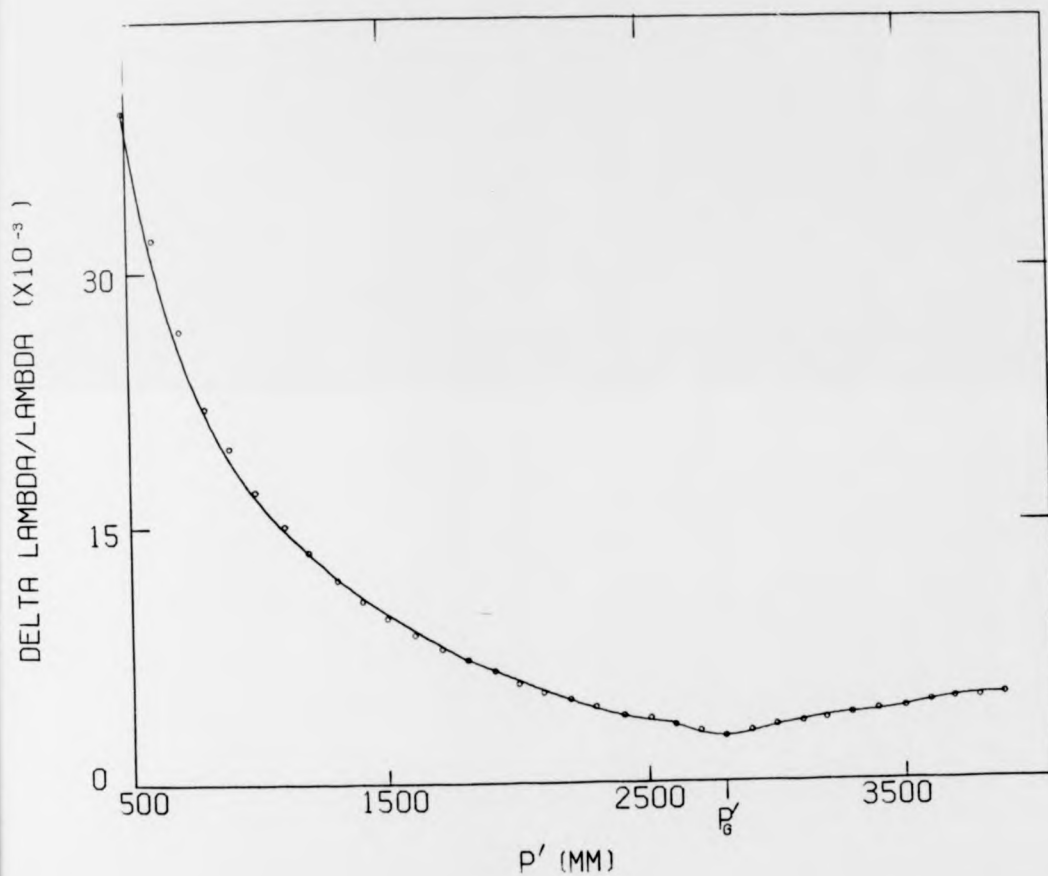


FIGURE 3.5 SPECTRAL RESOLUTION AS A FUNCTION OF P' FOR A GERMANIUM MONOCHROMATOR ON THE SRS (LAMBDA=1.488 ANGSTROMS)

oblique angle of incidence upon striking the film.

From studying the ray diagrams (figures 3.6 and 3.7) it can be seen that for the 'sample at focus' case:

$$\theta_x = \frac{\beta p' + s_x}{p' - l_x}$$

where β = intrinsic horizontal convergence of the X-ray beam;

s_x = collimator slit width.

For the 'detector at focus' case:

$$\theta_x = \frac{s_x}{p'} + \beta$$

Typical spot widths at the film for our data were of the order of 0.3 mm (crystal to film distance, 118 mm).

3.4.4 Fraction of the Available Synchrotron Beam Transmitted via the Collimator

For the 'sample at focus' case, this fraction is simply the ratio of the collimator exit slit width, s_x , and the width of the focus, v_x . The width of focus is the convolution of three factors, namely:

(1) The aberrations of the optical system which give a width of approximately:

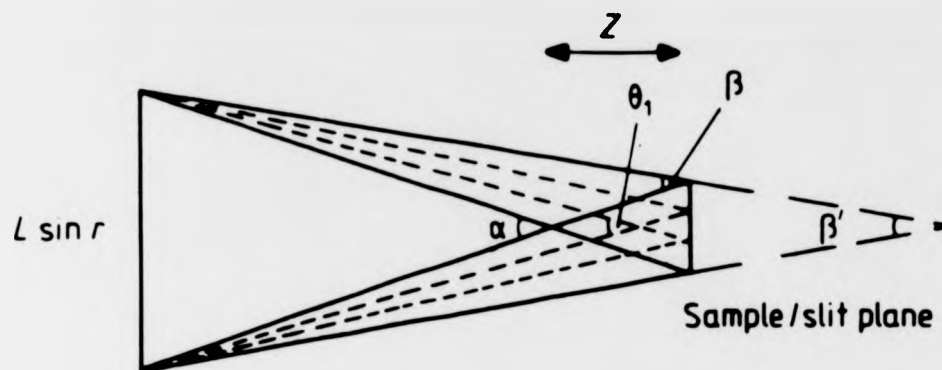


Figure 3.6 Ray Diagram Showing the Sample at the Focus

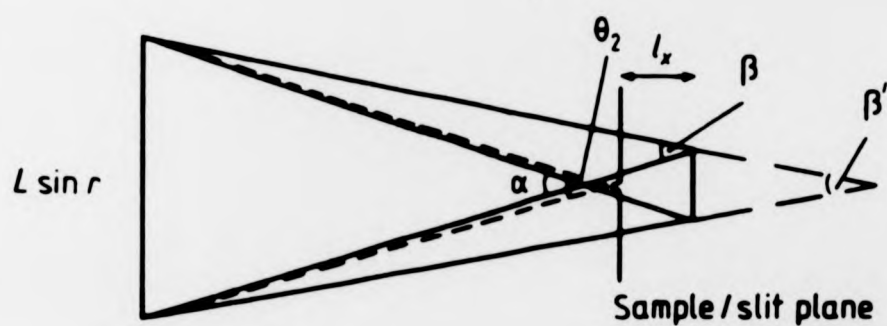


Figure 3.7 Ray Diagram Showing the Detector at the Focus

$$FC = \frac{L^2 \cos(\theta - \alpha)}{8R}$$

(see figure 3.10)

(Helliwell et al (1982))

- (2) The source demagnification, which gives a width of $\frac{hp'}{p}$.
- (3) The angular dispersion of the diffracted beam, the corresponding width is $\omega_s p'$

$$\text{where } \omega_s = 2.5 \times 10^{-4} |\cos 2\theta| \left(\frac{\sin r}{\sin i} \right)^{\frac{1}{2}}$$

(Lemonnier et al (1978)).

However, for the 'detector at focus' case the situation is a little more complex (see figure 3.7) and the fraction, f , is given by the formula:

$$f = \frac{s_x}{\beta' l_x + v_x}$$

where β' = convergence angle at focus

and v_x = horizontal width of beam at focus in the absence of a collimator.

This function has a maximum value of f at a particular p' given by:

$$(p')^2 = \frac{\beta_1 p^2 l_x}{h} \cdot \frac{\sin r}{\sin i}$$

where β_1 = acceptance angle from the ring

and h = horizontal width of incident beam.

$$FC = \frac{L^2 \cos(\theta - \alpha)}{8R}$$

(see figure 3.10)

(Helliwell et al (1982))

- (2) The source demagnification, which gives a width of $\frac{hp'}{p}$.
 (3) The angular dispersion of the diffracted beam, the corresponding width is $\omega_s p'$

$$\text{where } \omega_s = 2.5 \times 10^{-4} \left| \cos 2\theta \right| \left(\frac{\sin r}{\sin i} \right)^{\frac{1}{2}}$$

(Lemonnier et al (1978)).

However, for the 'detector at focus' case the situation is a little more complex (see figure 3.7) and the fraction, f , is given by the formula:

$$f = \frac{s_x}{\beta' l_x + v_x}$$

where β' = convergence angle at focus

and v_x = horizontal width of beam at focus in the absence of a collimator.

This function has a maximum value of f at a particular p' given by:

$$(p')^2 = \frac{\beta_i p_x^2}{h} \cdot \frac{\sin r}{\sin i}$$

where β_i = acceptance angle from the ring

and h = horizontal width of incident beam.

$$FC = \frac{L^2 \cos(\theta - \alpha)}{8R}$$

(see figure 3.10)

(Helliwell et al (1982))

- (2) The source demagnification, which gives a width of $\frac{hp'}{D}$.
- (3) The angular dispersion of the diffracted beam, the corresponding width is $\omega_s p'$

$$\text{where } \omega_s = 2.5 \times 10^{-4} |\cos 2\theta| \left(\frac{\sin r}{\sin i} \right)^{\frac{1}{2}}$$

(Lemonnier et al (1978)).

However, for the 'detector at focus' case the situation is a little more complex (see figure 3.7) and the fraction, f , is given by the formula:

$$f = \frac{s_x}{\beta' l_x + v_x}$$

where β' = convergence angle at focus

and v_x = horizontal width of beam at focus in the absence of a collimator.

This function has a maximum value of f at a particular p' given by:

$$(p')^2 = \frac{\beta_1 p_{1x}^2}{h} \cdot \frac{\sin r}{\sin i}$$

where β_1 = acceptance angle from the ring

and h = horizontal width of incident beam.

There is no such maximum for the 'sample at focus' case (see figure 3.8).

When setting up our experiments on the synchrotron workstation we wish to choose our monochromator to crystal distance so that we maintain a high flux at the sample (in order to minimise exposure times), while at the same time wishing to keep the size of the diffracted spots small. By keeping the spots small we increase their optical density and reduce the risk of distinct spots overlapping. Therefore the ratio of these two, $\frac{f}{w_H}$ is a useful parameter to calculate when deciding where to place our crystal. Graphs showing the variation of $\left[\frac{f}{w_H}\right]_{\max}$ with sample size, mosaic spread, and beam width for both the sample at focus and detector at focus cases are shown in figure 3.9. The form of these graphs is the same for both cases, however, the value of $\left[\frac{f}{w_H}\right]_{\max}$ is always lower with the detector at the focus. It is therefore apparent that to have the sample at the focus is preferable. This is also consistent with the simple procedure of alignment of the camera to the focussed synchrotron beam by maximising the flux through the collimator (see section 3.5.6). If focussing the beam at the detector had been optimal then one would have required a fine pixel size electronic area detector (such as a charged coupled device (Allinson (1982)) in order to optimise the camera alignment.

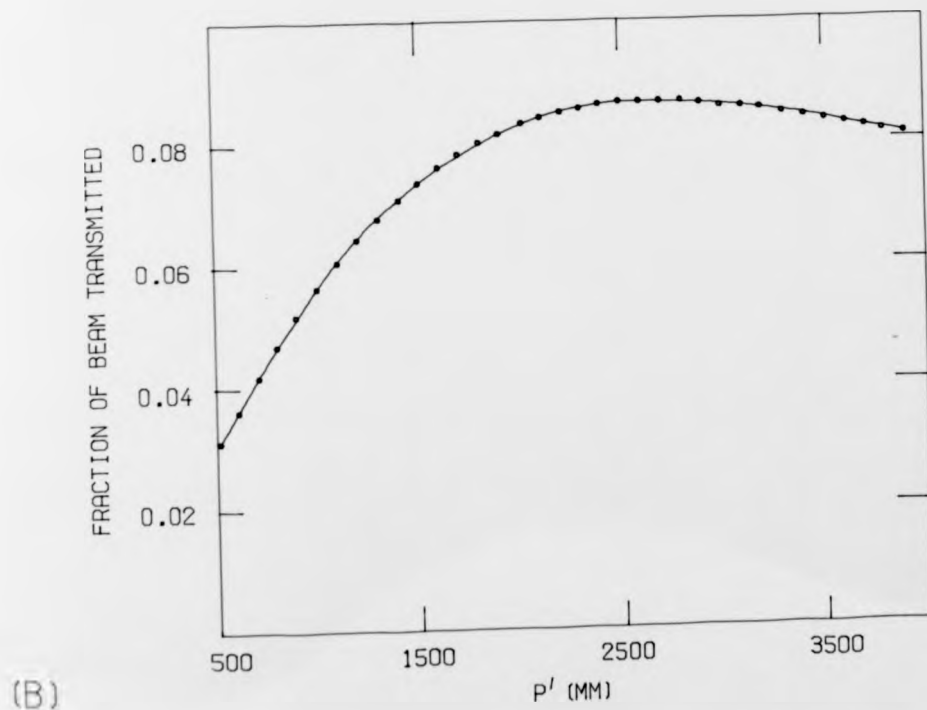
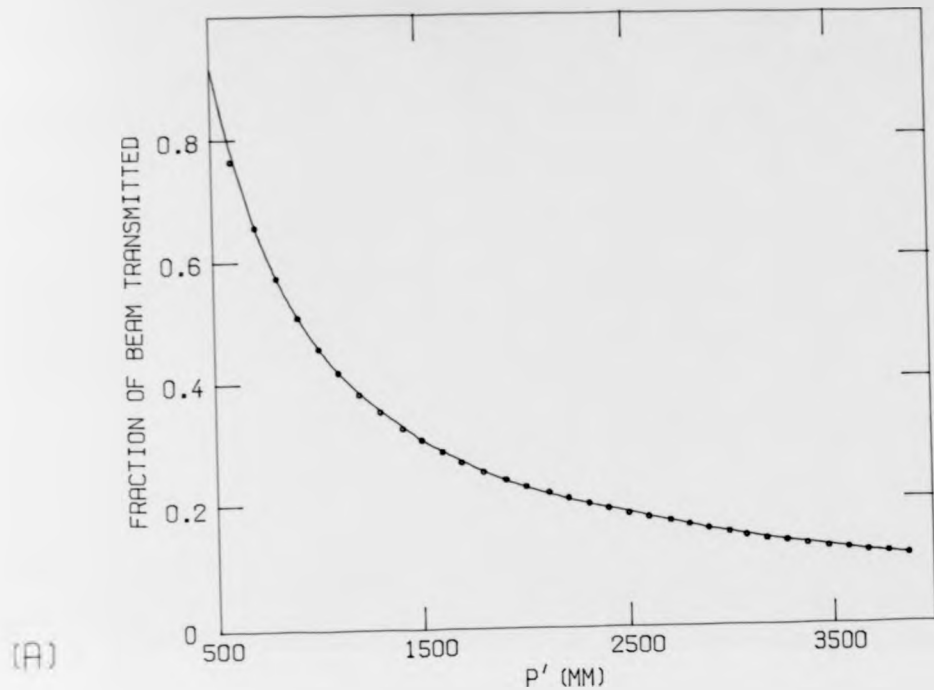


FIGURE 3.8 GRAPHS SHOWING THE VARIATION OF THE FRACTION OF AVAILABLE SYNCHROTRON BEAM TRANSMITTED FOR (A) SAMPLE AT FOCUS AND (B) DETECTOR AT FOCUS

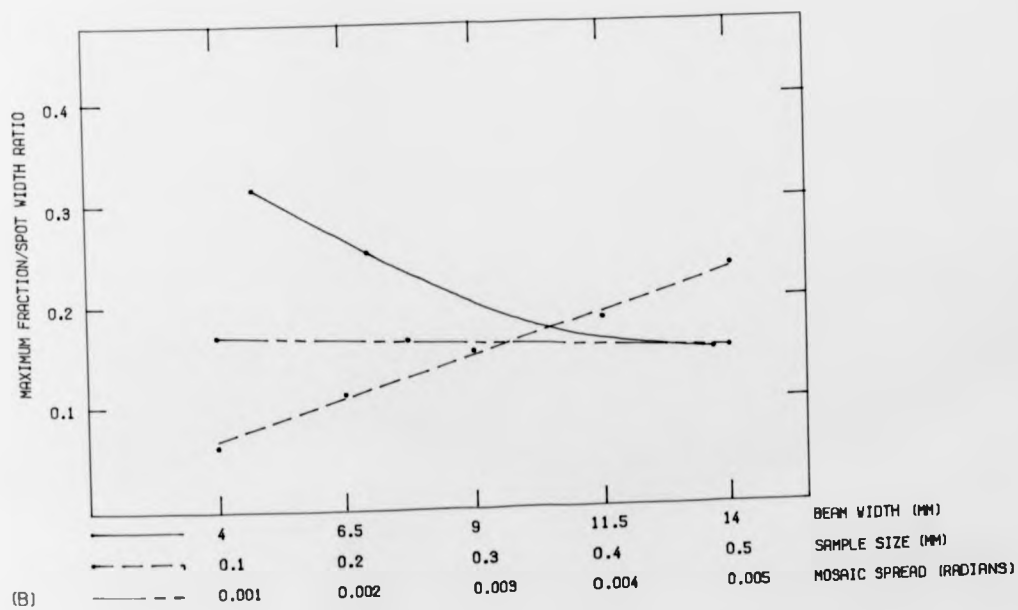
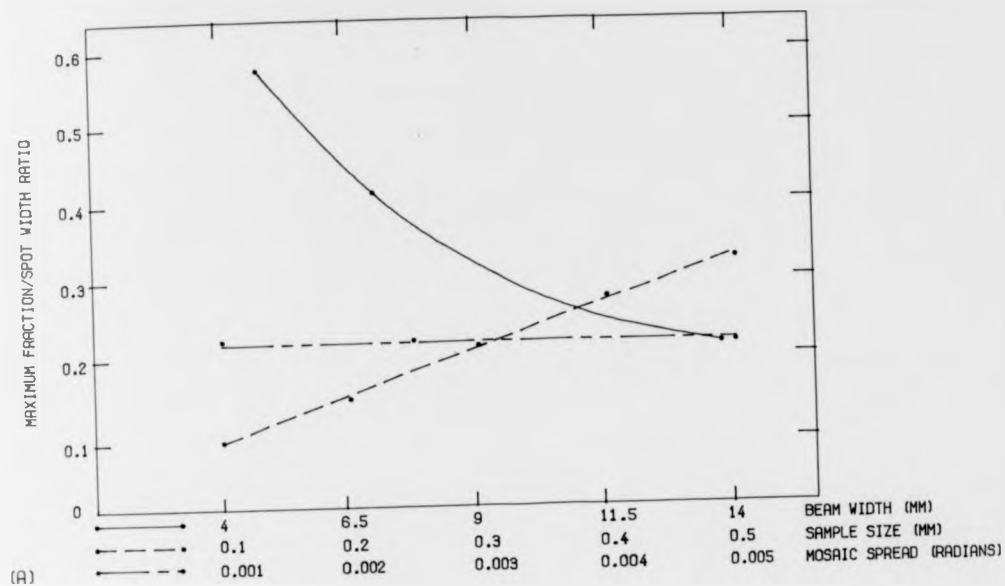


FIGURE 3.9 GRAPHS SHOWING THE VARIATION OF THE RATIO OF THE FRACTION OF AVAILABLE BEAM TO HORIZONTAL SPOT WIDTH FOR (A) SAMPLE AT FOCUS AND (B) DETECTOR AT FOCUS

3.4.5 Depth of Focus

It would be of little use taking great care in placing the sample in the optimal position according to the calculations outlined above if the system had a very large depth of focus (of the order of p' , say). By studying the Rowland Geometry for the system (see figure 3.10), it can be seen that there is effectively a zero 'depth of focus' and a definite minimum width is clearly shown at the point CF. There is, however, some lengthening of this focus due to the 'depth of source' of the system. Photons originate from different points along the electron orbit and are therefore focussed to different distances from the monochromator. For the SRS with a bending radius of 5.5 m, if 4 milliradians of beam are accepted then this depth of source is 22 millimetres and the corresponding lengthening of the focus is 2.64 millimetres allowing for the demagnification of the source depth by the asymmetry factor A (see section 3.4). This figure was calculated for the germanium monochromator (oblique cut 10.5°) at the Guinier position for $\lambda = 1.488 \text{ \AA}$.

3.5 The Experimental Workstation

The protein crystallography workstation on which my data was collected and with which I have been involved with at the design and commissioning stage is situated on X-ray beam line 7 of the SRS. The synchrotron radiation emitted

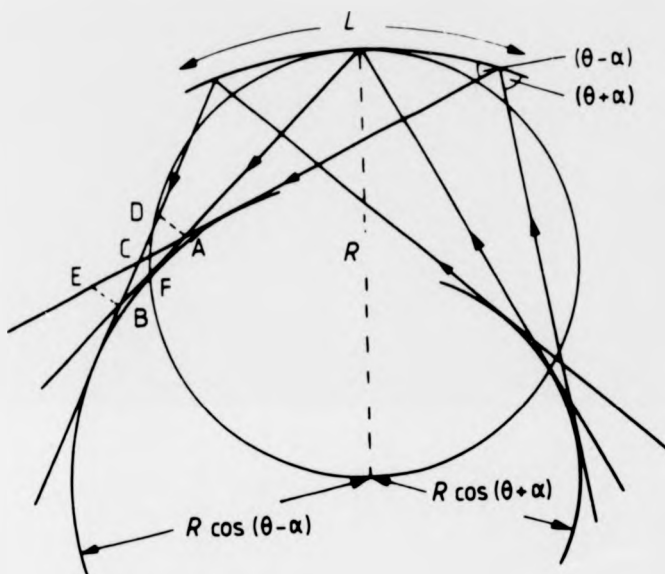


Figure 3.10 Rowland Geometry Showing Focus at CF

from the ring enters the evacuated stainless steel beam line via a tangent port immediately following bending magnet 7. Twenty-eight milliradians of beam enter the port which is then split into three usable sections of four milliradians, one to be used for protein crystallography and small angle scattering experiments, the second carries on into the outer hall and is used by interferometry and topography experiments, the final four milliradians is used for EXAFS experiments.

After leaving the beam splitter the beam is reflected from a flat, platinum coated, fused, quartz mirror. The mirror has a couple applied to each end which produces a cylindrical surface with a bending radius of approximately 1 km. This mirror focusses the beam in the vertical direction but leaves the horizontal divergence (4 milliradians) unaltered. The mirror is located 11 m from the tangent port (see fig. 3.11).

The beam then travels a further 10 m to the monochromator (see section 3.5.2). The monochromator vessel has slits both before and after the perfect crystal to control the amount of white radiation illuminating the crystal and also the size of monochromatic beam reflected (see section 3.4). The vacuum pressure is maintained at approximately 5×10^{-8} torr between the tangent port and a Be window just prior to the monochromator preslits.

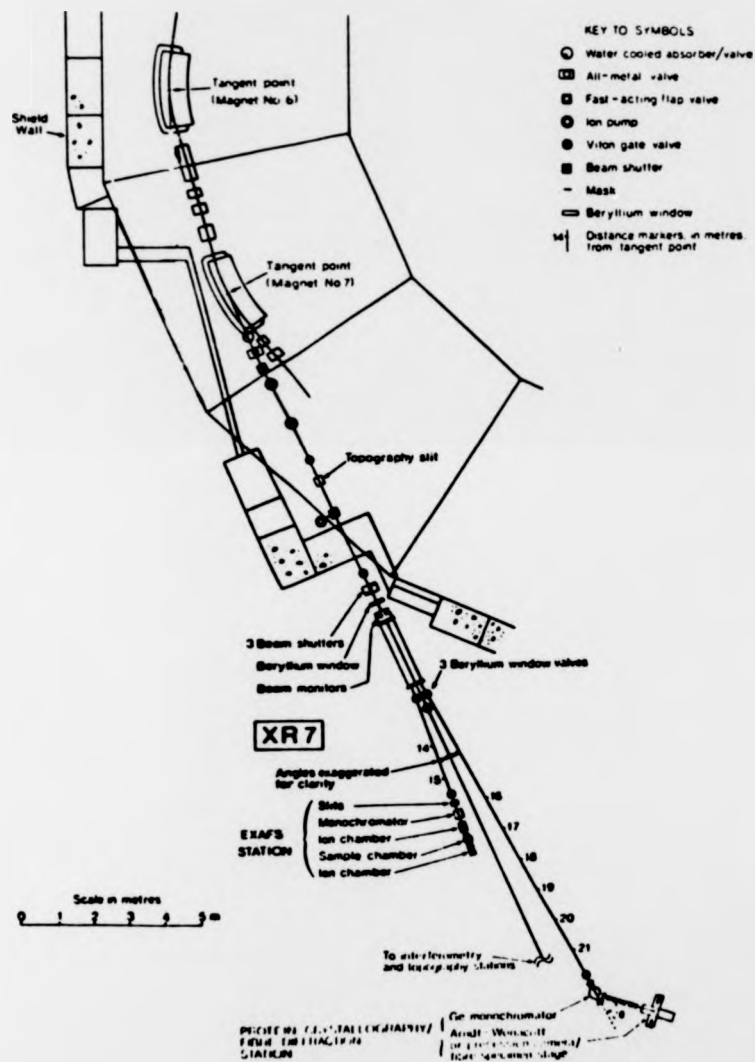


Figure 3.11 Scaled Diagrammatic Representation of X-Ray Beam Line Seven

The rest of the apparatus beyond the monochromator vessel is supported upon an adjustable experimental arm (see section 3.5.3). Between the Beryllium window at the exit of the monochromator vessel and the camera there is an evacuated steel pipe with low absorption 'kapton' windows $10\mu\text{m}$ thick at each end. The vacuum pressure in the monochromator vessel and this last section of steel pipe is approximately one torr. Beyond the steel pipe is a modified Arndt-Wonacott rotation camera supported by a remotely controlled alignment carriage (see section 3.5.4).

3.5.1 Slit design

Both prior to and following the monochromator there are a series of slits. Each unit consists of four independently operated jaws which are powered by Astrosyn stepper motors via a dedicated PDP11/04 computer control system (see section 3.5.5). these slits are made of a tungsten/lead/aluminium composite with the tungsten biting into the synchrotron beam and the lead and aluminium absorbing the resulting fluorescent radiation. The aluminium is machine cut and bevelled inwards by 3° to minimise scatter from the slit edges and thus reduce background scatter (see fig. 3.12).

The premonochromator slits are used to prevent synchrotron radiation falling on the clamp and also to restrict the length of the monochromator that is illuminated

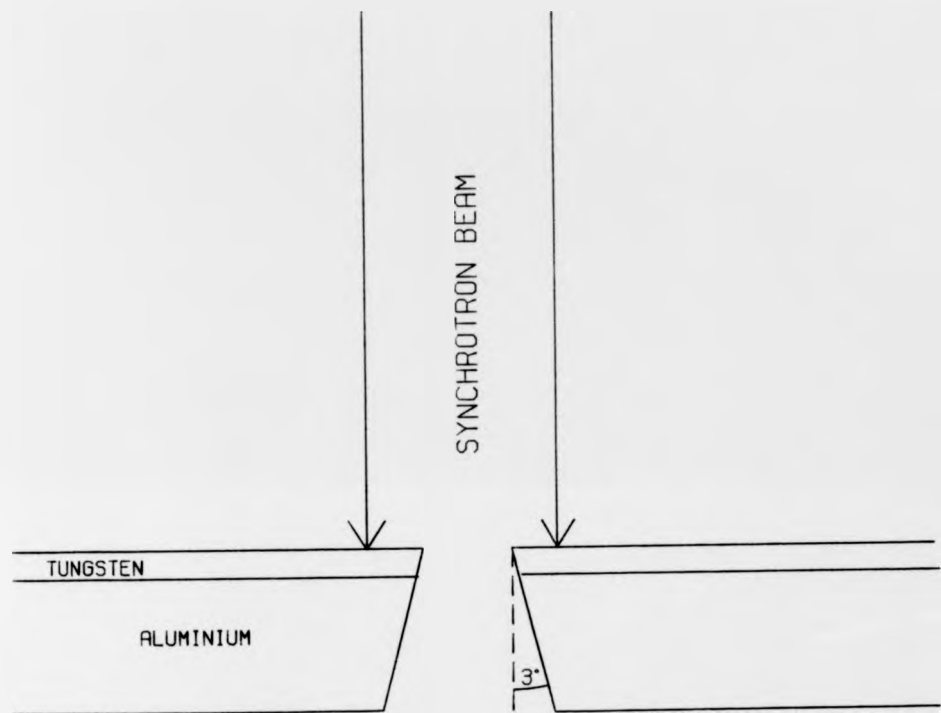


FIGURE 3.12 DIAGRAM SHOWING THE SLIT DESIGN

as well as controlling the horizontal and vertical beam convergence (and therefore the size of Bragg reflections (see section 3.4.3)). The postmonochromator slits are used to ensure that a suitable amount of monochromatic beam is directed towards the collimator and that background scatter from the various optical elements are reduced. This reduction of parasitic scatter is of importance for small angle scattering experiments (SAXS).

3.5.2 The Monochromator Vessel

The perfect crystal triangular monochromator is clamped at its base with its apex resting against an eccentric cam. The use of the eccentric cam to bend the crystal means that it cannot be cracked by accidental overstressing. The whole assembly is mounted on a goniometric system and rotary table (see fig. 3.14). This allows six degrees of freedom (3 translation and 3 rotation). The system is powered by Astrosyn stepper motors and controlled by the PDP11/04 computer control system (see section 3.5.5). The apparatus is housed in an evacuated vessel (vacuum pressure approximately 1 torr) with an exit window made of 0.5 mm thick Berillium curved into a 90° arc. This window is 20 mm wide.

3.5.3 The Experimental Arm

The experimental arm is a stable four metre bench which

is supported by three thrust bearings in a unique design (P.R. Moore, pers. comm.) see figures 3.13 and 3.14. One bearing is directly below the monochromator vessel at the point where the experimental arm pivots while the other two support the experimental arm by means of bushes connecting the upper half of each bearing to a separate rail running parallel below the bench along approximately half of its length. The lower half of each thrust bearing is connected, again via bushes, to horizontal cross track rails that are permanently attached to the floor. These cross tracks are at an angle of approximately 55° to the direct beam (actually surveyed as 0.9593 radians).

The result of this construction is a large stable bench that can easily be moved by hand through a 2θ range of seventy degrees. Once in position the arm is easily clamped via a metal plate connected to the lower half of one of the bearings.

A recent (September 1983) refinement to the system is the addition of a motorised unit which allows accurate 'fine tune' movements over a range of 400 mm. The required 2θ angle is selected manually by moving the experimental arm but instead of clamping the bearing one clamps the motorised unit in place which allows one to operate the arm via the motor control keyboard (see section 3.5.5).

The motorised unit (see figure 3.15) consists of a

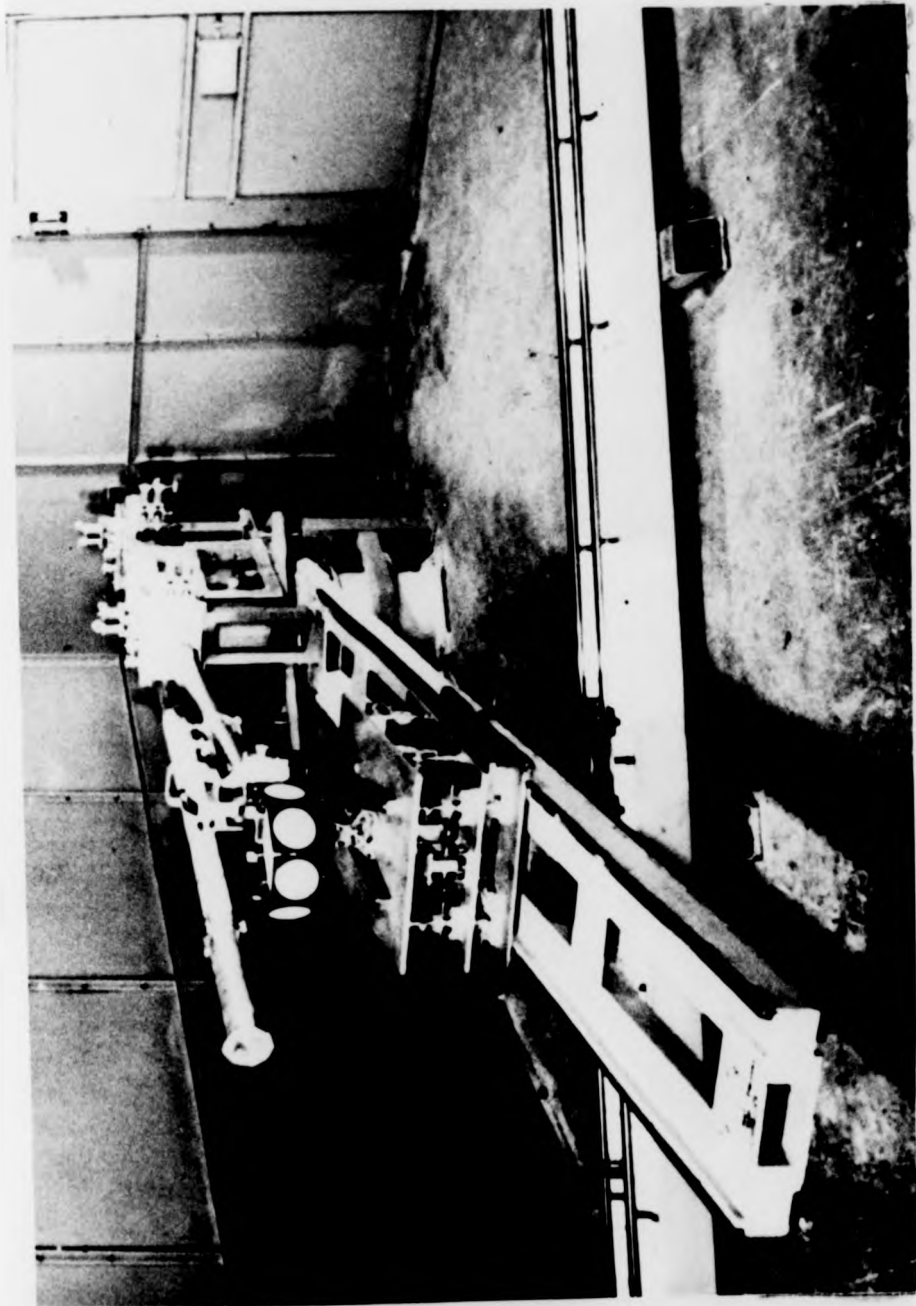


Figure 3.13 Photograph of the Workstation showing
the Experimental Arm

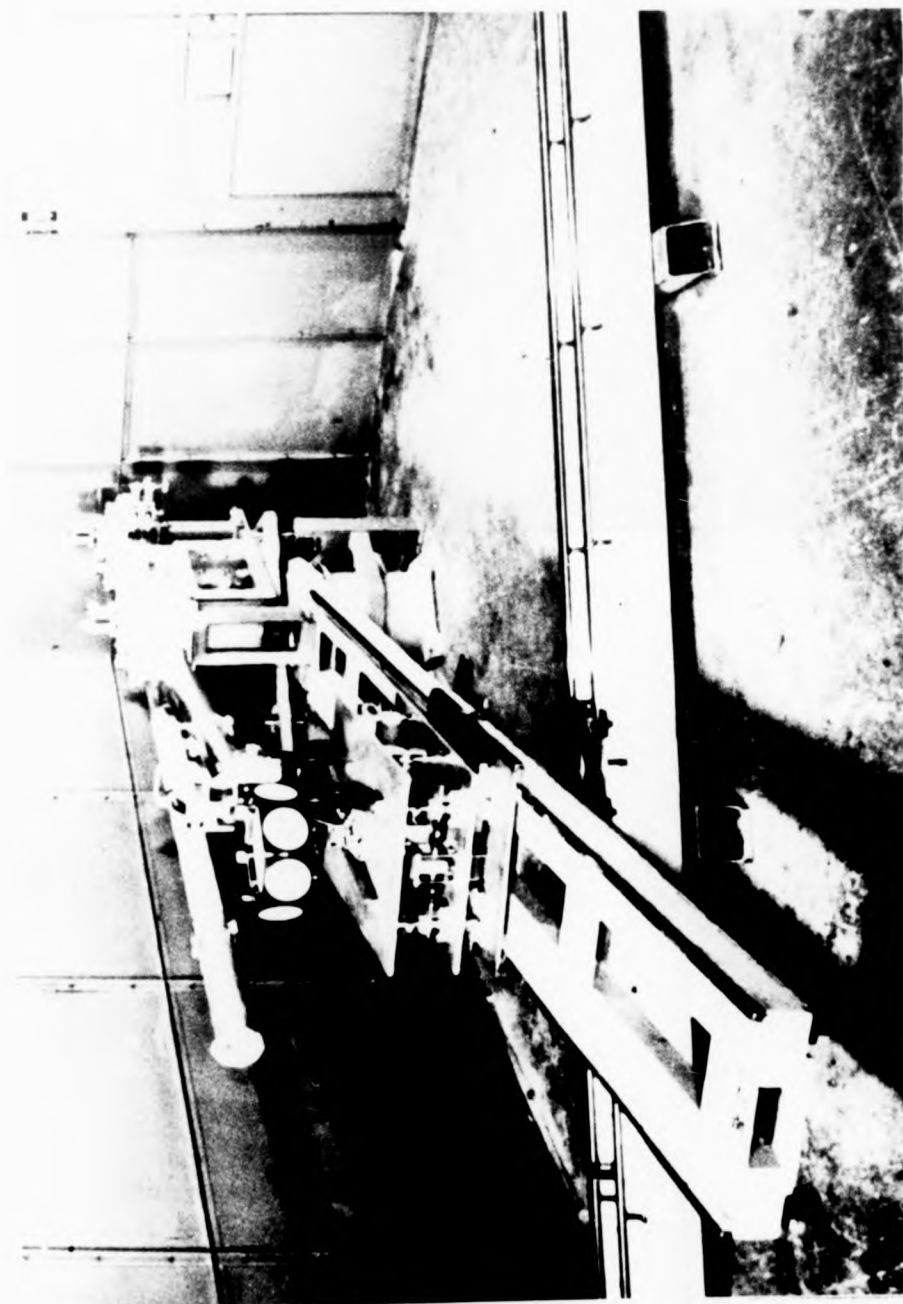


Figure 3.13 Photograph of the Workstation showing
the Experimental Arm

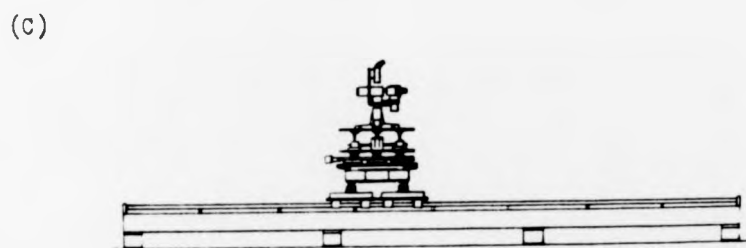
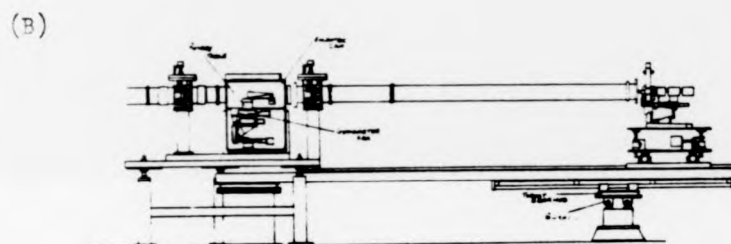
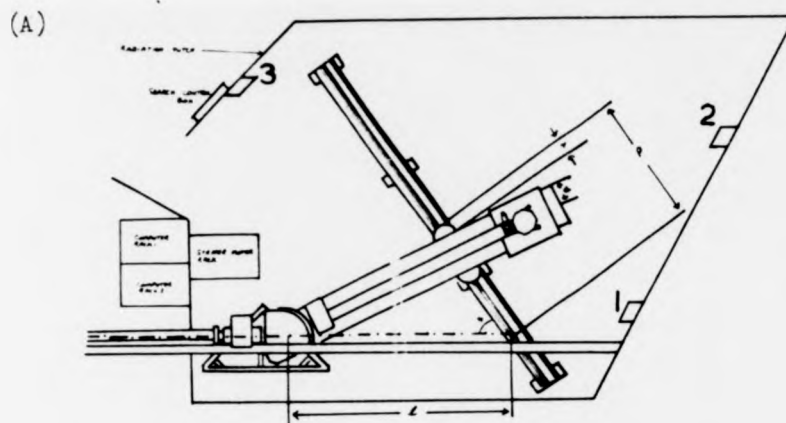


Figure 3.14 Draftsman's Drawings of the Experimental Arm
(A) Plan (B) Side Elevation (C) Front Elevation (Based on
Daresbury Laboratory Drawings 60/177 and 60/183)

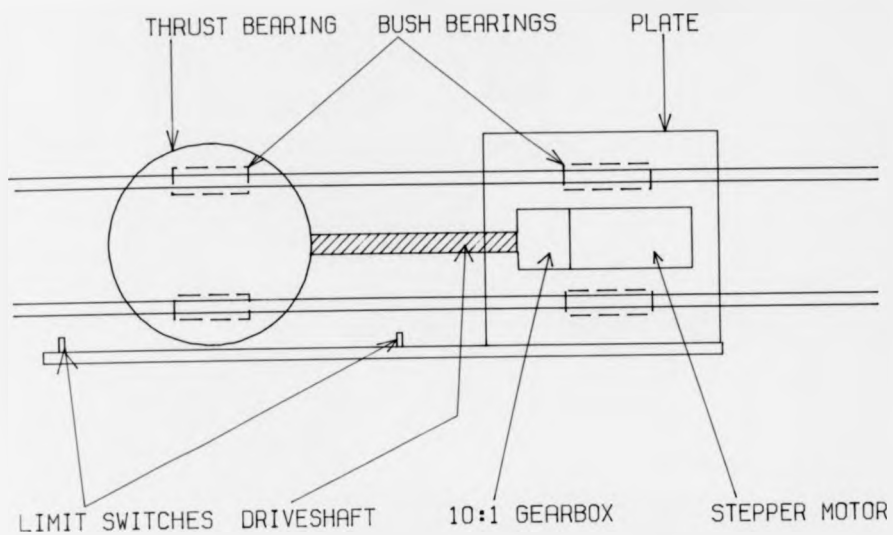


FIGURE 3.15 SKETCH SHOWING THE PRINCIPLE FEATURES OF THE BRAGG
ANGLE CONTROL DRIVER

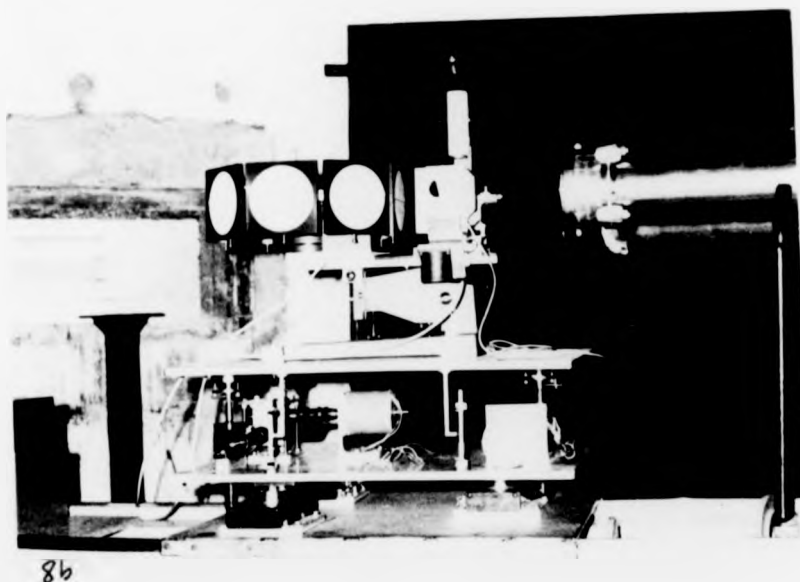


FIGURE 3.16 PHOTOGRAPH OF THE ALIGNMENT CARRIAGE

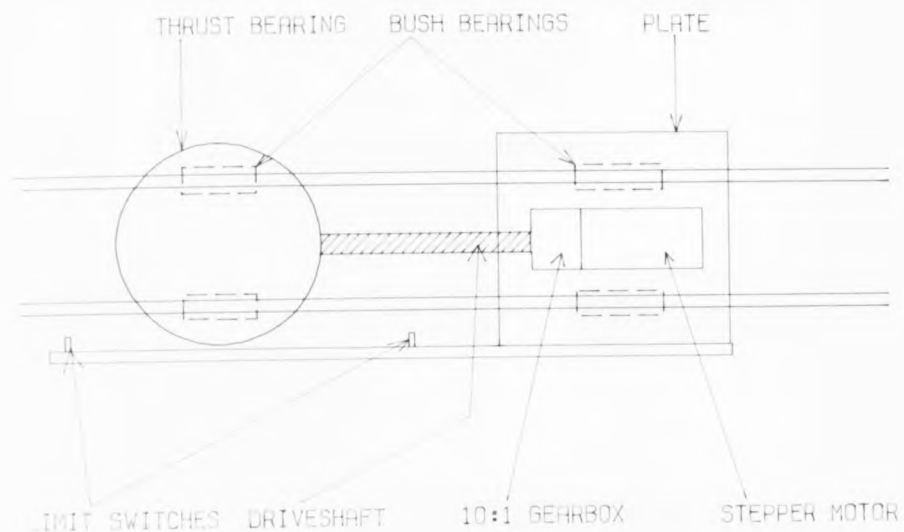
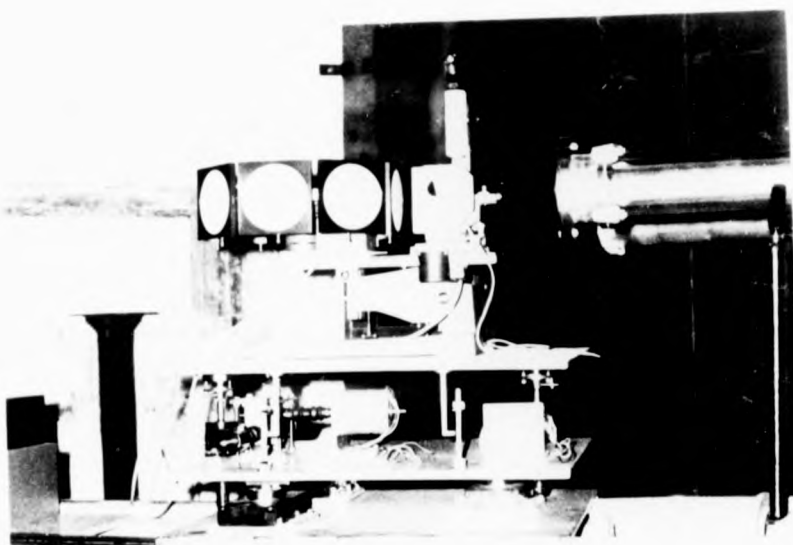


FIGURE 3.15 SKETCH SHOWING THE PRINCIPLE FEATURES OF THE BRAGG
WHEEL CONTROL DRIVER



86

FIGURE 3.16 PHOTOGRAPH OF THE ALIGNMENT CARRIAGE

metal plate supported by bushes onto the cross tracks which is connected to the lower half of one of the thrust bearings via a ball screw and ball nut. The ball screw is driven by a stepper motor connected via a 10:1 gearbox. Both the stepper motor and gearbox are bolted to the metal plate. A single step of the motor results in a $2.5\mu\text{m}$ movement along the track which corresponds to a change of 2θ angle of 6×10^{-5} degrees for $\lambda = 1.488$. This is much less than the smallest $\delta 2\theta$ obtained by movement of the monochromator rotary table ($\delta 2\theta = 0.001^\circ$)

One advantage of having motorised movement of the experimental arm is that changes in the 2θ angle can be simply made. Previously the alignment carriage had to be subjected to a horizontal translation followed by a horizontal rotation to compensate for any monochromator movements (see section 3.5.6) during wavelength scanning.

In order to set the experimental arm to the required wavelength we need a relationship between the distance along the cross track of the experimental arm and the 2θ angle of the monochromator. If we use the right hand edge (looking towards the monochromator) of the right hand bearing as a pointer, I calculate that the distance of this from the point of intersection of the direct beam flight path and the cross tracks (distance P on figure 3.14) is related to 2θ by the following:

$$P = \frac{l \sin 2\theta}{\sin(2\theta + \alpha)} + \frac{d}{(\sin 2\theta + \alpha)} + r$$

where l = direct beam monochromator to cross track distance
 = 3052.33 mm

α = angle between direct beam and cross track = 0.9593
 radians

d = half width of experimental arm = 206 mm

r = radius of thrust bearing = 200 mm

θ = Bragg angle

3.5.4 The X-ray Camera and Alignment Carriage

The camera used on the protein crystallography workstation at the SRS is an Arndt-Wonacott Rotation Camera which is fully described by Arndt et al (1973). The camera is mounted on top of an alignment carriage specially built by the University of Keele Physics Department workshops. (Dudley, Greasley and Marsh pers. comm.). The carriage has four degrees of freedom of movement (two rotation and two translation about axes perpendicular to the direct beam), figure 3.16. The table top dimensions are 600 mm x 600 mm, with rotation axes being very close to the collimator entrance hole. The design is based on kinematic principles, capable of 5 μ m translations or angular tilts of 10^{-5} radians (5 μ m over 600 mm). Total movements are controlled by limit switches and are approximately ± 25 mm. The carriage moves on recirculating roller bearings along the experimental arm to allow the focussing distance of the monochromator to be adjusted over a range of $p' = 0.7$ to 3.5 m from the monochromator. This allows selection of the Guinier

condition for a wide range of wavelengths.

3.5.5 The Computer Control System

The control of the workstations 22 stepper motors (8 for slit control, 7 for monochromator control, 4 for alignment carriage control, 2 for mirror control and 1 for experimental arm (20) control) is performed by a dedicated PDP11/04 computer via a CAMAC based motor address system. A schematic flow chart of the system is given in figure 3.17.

The control software resident in the PDP11/04 memory receives information from the user via a portable keyboard as to which motor(s) are required to be moved and what speed and in which direction and also whether a single step or continuous motion of the motor is required. The software then transmits the appropriate octal addresses to the stepper motor multiplexers which are in turn connected 'daisy chain' fashion to the stepper motors via Highway Interface Modules (HIM's), see figure 3.17. These HIM's interface the stepper motor power supplies with the controlling multiplexers and amplifiers. The software also keeps track of all motor positions and displays these on the VDU screen upon request from the VDU keyboard. It also displays individual motor positions whenever they are moved.

The original software which was used up until January 1982 was written in a BASIC-related language called CATY

(Greenhough T.J. pers. comm.). It did not have the use of a portable handheld keyboard but instead had a large static keyboard. The mark II software, which was written by myself in a different but related language (CATEX) and had several novel features in addition to being written in a format that is far easier to understand and therefore to edit. The novel features include:

- (1) Written in CATEX;
- (2) Single stepping of any motor allowed;
- (3) Utilisation of a pocket calculator type keyboard with alphanumeric cursor display.
- (4) All motor positions can be automatically dumped to a data file.

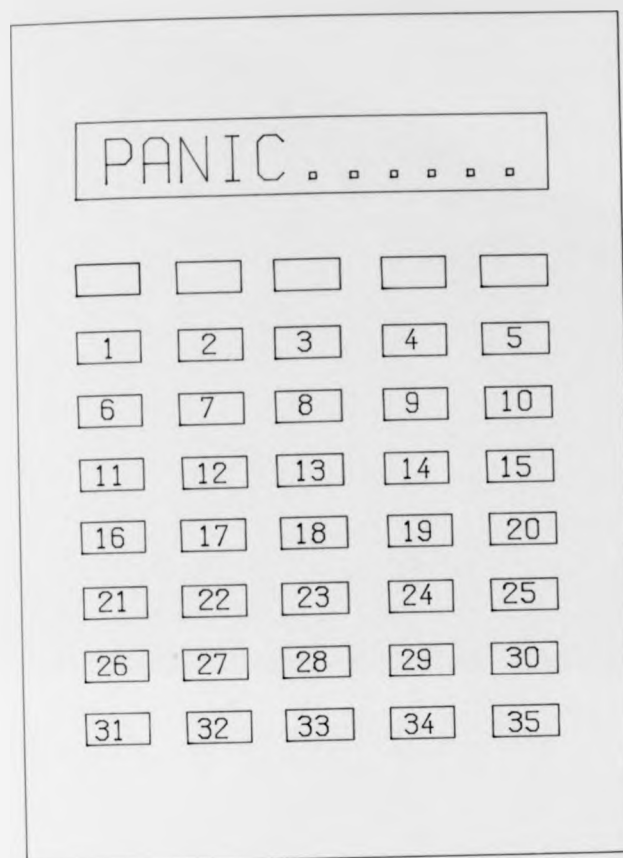
By using CATEX the code can be written in a far more understandable way. This is because CATEX allows variable and subroutine names to be alphanumeric strings rather than single letters. CATEX also allows the use of the FORTRAN CALL command to invoke named subroutines as well as the BASIC GOSUB type of subroutines.

The single step option allows very small and accurate ($5\mu\text{m}$) movements to be made. This is very useful especially when mirror or monochromator adjustments are being made. The previous system relied on the speed and skill of the user to switch the motors on and off again quickly when making fine adjustments.

The use of a pocket calculator type keyboard with integral alphanumeric cursor display allows adjustments to be made both inside and outside the experimental 'hutch' (see section 3.5.7). the printing out of words such as FAST and SLOW or S.STEP and M.STEP on the cursor is less confusing than the LED type of display present on the earlier system. The pocket terminal and key assignments are shown in figure 3.18).

A further improvement on the mark I software was the ability of the system to dump all of the current motor positions to a data file called 'info.dat' when returning to the RT11 operating system. This means that if the computer is required for program development, or other uses, it is easy to reload the stepper motor software just as it was.

A further important feature of the software that was also available in the mark I version is the presence of a 'panic stop' button (see figure 3.18). Once pressed this button causes all of the motors to stop and then disables them all. No motor may subsequently be moved unless the panic stop is reset. By using this option once all motors are in their optimal positions one eliminates the possibility of accidental movement of any motor during data collection due to the user inadvertently pressing a button on the keyboard. Minor modifications to my original software have been made by A.W. Thompson (pers. comm.) to include the experimental arm (2θ) motor into the package.



← CURSOR CONTROL

- 1 PRE SLITS (V.BOTTOM)
- 2 PRE SLITS (V.TOP)
- 3 PRE SLITS (H.INNER)
- 4 PRE SLITS (H.OUTER)
- 5 POST SLITS (V.BOTTOM)
- 6 POST SLITS (V.TOP)
- 7 POST SLITS (H.OUTER)
- 8 POST SLITS (H.INNER)
- 9 MONOCHROMATOR (X-TRANS)
- 10 MONOCHROMATOR (Y-TRANS)
- 11 MONOCHROMATOR (Z-TRANS)
- 12 MONOCHROMATOR (ROT. TABLE)
- 13 MONOCHROMATOR (THETA -X)
- 14 MONOCHROMATOR (THETA -Y)
- 15 MONOCHROMATOR CAM
- 16 EXP. ARM (TWO THETA)

- 17 ALIGNMENT CARRAGE (H.FRONT)
- 18 ALIGNMENT CARRAGE (H.BACK)
- 19 ALIGNMENT CARRAGE (V.BACK)
- 20 ALIGNMENT CARRAGE (V.FRONT)
- 21 ALIGNMENT CARRAGE (H.TRANSLATION)
- 22 ALIGNMENT CARRAGE (H.ROTATION)
- 23 ALIGNMENT CARRAGE (V.TRANSLATION)
- 24 ALIGNMENT CARRAGE (V.ROTATION)
- 25 MIRROR
- 26 MIRROR
- 27-30 UNASSIGNED
- 31 ABORT CATX
- 32 SINGLE/MULTIPLE STEP
- 33 MOTOR DIRECTION (CLW/ACLW)
- 34 MOTOR SPEED (FAST/SLOW)
- 35 PANIC STOP

FIGURE 3.18 SKETCH SHOWING POCKET TERMINAL AND KEY ASSIGNMENTS

A data acquisition package that allows exposure times to be varied as a function of ion chamber (flux) readings along with automatic rotation of the camera spindle and carousel, as suggested by Bartunik et al (1981), is currently being written (Sept 1983 - A.W. Thompson, pers. comm.). This package will supplement the mark II software but unfortunately due to lack of core cannot be loaded simultaneously. Although this new package will allow adjustment to be made to the alignment carriage it cannot be used to adjust the monochromator or slit motors.

Microfiches of the mark I, mark II and new data acquisition packages can be found in Appendix I.

3.5.6 Alignment Procedure

Most of the alignment procedure and focussing of the X-ray beam is done by means of maximising the ionisation current produced by an ionisation chamber made up of an ebony horseshoe with pieces of copper clad circuit board forming the plates (see figure 3.19). The chamber is operated at a potential of -100V. The signal received on the second plate is collected by a Keithley 427 current amplifier with the gain normally set at 10^9 (i.e. 1 volt output for 1 nanoamp of ionisation current detected). This voltage is then displayed on a digital voltmeter. The detector shows good linearity between output current and incident photon flux despite the simple design.



Figure 3.19 Photograph Showing the Air Ionisation Chamber
Used to Monitor Beam Flux

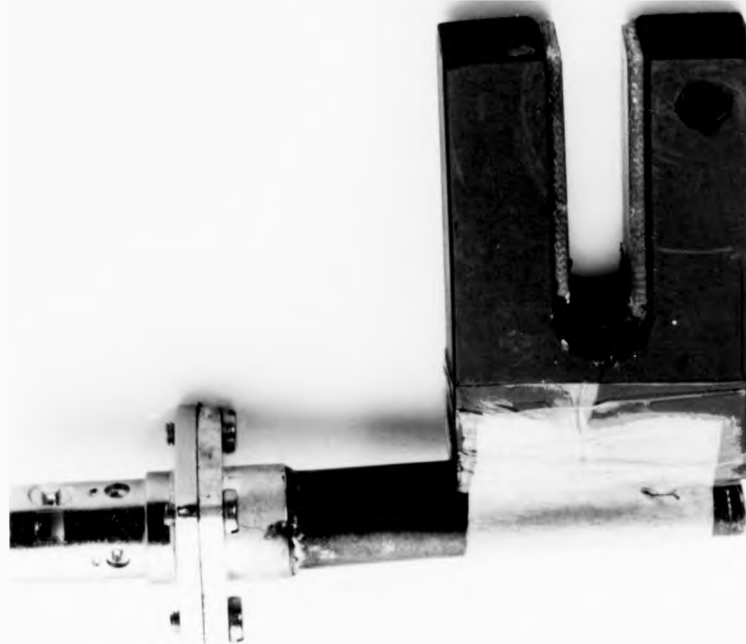


Figure 3.19 Photograph Showing the Air Ionisation Chamber
Used to Monitor Beam Flux

The first step in alignment is the selection of the required wavelength. The most accurate method available requires setting the wavelength to known absorption edges (which are usually the wavelengths of practical interest). A list of absorption edges are given in table 3.1. The experimental arm is first clamped into position at the required 2θ angle, as calculated using the formula given in section 3.5.3. The monochromator is then rotated through an angle calculated to bring it slightly short of the required position. The collimator is then aligned to the beam by adjusting the alignment carriage motors and the monochromator cam (see later for alignment procedure) until a maximum reading is obtained on the ion chamber output display. This is recorded and then the monochromator is rotated a further single step ($0.001^\circ \equiv 0.5 \text{ eV}$ at 1.488 \AA). The alignment carriage and monochromator cam are again adjusted for maximum flux which is then recorded. This procedure is repeated until a full transmission spectrum of the absorption edge is plotted (see figure 3.20). The monochromator is then rotated back to the edge position. This procedure is usually carried out by the stationmaster (J. Helliwell pers. comm.). The addition of the experimental arm motor will simplify this procedure because once the collimator has been aligned to the beam the first time, subsequent changes in 2θ angle can be adjusted for by simply translating the experimental arm. No further misalignment should have occurred as the monochromator rotary table is set horizontally.

TABLE 3.1
Wavelengths of K-emission Series and Absorption Edges
The units are kX 1.00202 They may be taken as in Å to an accuracy of 1 in 25,000 only (see Section 2.1.2).

At. No.	Element symbol	A_{α_2}	A_{α_1}	$A\beta_1$	$A\beta_2$ (D signifies mean of doublet)	A absorption edge
1	H					
2	He					226.5
3	Li	240				
4	Be	113				
5	B	67				
6	C	44				43.5 _{av}
7	N	31.60				30.5 _{av}
8	O	23.71				23.2 _{av}
9	F	18.31				
10	Ne	14.616		14.464		
11	Na	11.909		11.617		9.5117
12	Mg	9.8889		9.558		7.9511
13	Al	8.33916	8.33669	7.981		6.7446
14	Si	7.12773	7.12528	6.7681		5.7866
15	P	6.1549		5.8038		
16	S	5.37471	5.37196	5.03169		5.0182
17	Cl	4.73050	4.72761	4.4031		4.3969
18	Ar	4.19456	4.19162	3.8848*		3.8707
19	K	3.74462	3.74122	3.5538		3.43645
20	Ca	3.36159	3.35825	3.0896		3.07016
21	Sc	3.03452	3.03114	2.7795		2.757 ₁
22	Ti	2.75207	2.74841	2.51381		2.497 ₁₀
23	V	2.50729	2.50348	2.28434		2.269 ₁₀
24	Cr	2.29351	2.28962	2.08480		2.070 ₁₇
25	Mn	2.10568	2.10175	1.91015		1.896 ₂₀
26	Fe	1.93991	1.93597	1.75653		1.743 ₁₁
27	Co	1.79278	1.78892	1.62075	1.48861	1.606 ₁₁
28	Ni	1.66169	1.65784	1.50010	1.38102	1.488 ₁₂
29	Cu	1.54433	1.54051	1.39217*	1.28366	1.380 ₁₁
30	Zn	1.43894	1.43511	1.29522		1.283 ₁
31	Ga	1.34394	1.34003	1.20784	1.19595	1.195 ₁
32	Ge	1.25797	1.25401	1.12890	1.11682	1.116 ₁
33	As	1.17981	1.17581	1.05726	1.04498	1.044 ₁
34	Se	1.10875	1.10471	0.99212	0.97986	0.979 ₁
35	Br	1.04376	1.03969	0.93273	0.92084	0.91995
36	Kr	0.9841	0.9801	0.87845	0.86609	0.86547
37	Rb	0.92963	0.92551	0.82863	0.81641	0.81549
38	Sr	0.87938	0.875214	0.78288	0.77076	0.76969
39	Y	0.83401	0.82879	0.74068	0.72874	0.72762
40	Zr	0.79010	0.78588	0.70170	0.68989	0.68877
41	Nb	0.75040	0.74615	0.66572	0.65412	0.65291
42	Mo	0.713543	0.70926	0.63253	0.62099 (D)	0.61977
43	Tc	0.67927*	0.67493*	0.60141*	0.59018*	0.58911
44	Ru	0.64736	0.64304	0.57246	0.56164	0.56017
45	Rh	0.617610	0.613245	0.54559	0.53509 (D)	0.533 ₁₀
46	Pd	0.589801	0.585415	0.52052	0.51021	0.5091 ₁₀
47	Ag	0.563775	0.559363	0.49701	0.48701	0.4858 ₁₀
48	Cd	0.53941	0.53498	0.475078	0.46531	0.46409
49	In	0.51652	0.51209	0.454514	0.444963	0.44388
50	Sn	0.49502	0.49056	0.435216	0.425900	0.42468
51	Sb	0.47479	0.470322	0.417060	0.407950	0.40663
52	Te	0.455751	0.451263	0.399972	0.391080	0.38972
53	I	0.437805	0.433293	0.383884	0.37547	0.373 ₁₀
54	X	0.42043	0.41596	0.36846	0.35989	0.35849
55	Cs	0.404812	0.400268	0.354347	0.346084	0.34474
56	Ba	0.389646	0.385089	0.340789	0.332745	0.33137
57	La	0.375729	0.370709	0.327959	0.32024 (D)	0.31842
58	Ce	0.361665	0.357075	0.315792	0.30826 (D)	0.30647
59	Pr	0.348728	0.344122	0.304238	0.29690 (D)	0.29516
60	Nd	0.336487	0.331822	0.293274	0.28631	0.28451
61	Pm	0.3249	0.3207	0.28209	(0.2761)	(0.2743)
62	Sm	0.31365	0.30895	0.27305	0.26629	0.26462
63	Eu	0.30326	0.29850	0.26360	0.25697	0.25552
64	Gd	0.29320	0.28840	0.25445	0.24812	0.24680
65	Tb	0.28343	0.27876	0.24601	0.23960	0.23840
66	Dy	0.27430	0.26957	0.23758	0.23175	0.23046
67	Ho	0.26552	0.26083	(0.2302)	(0.2244)	0.22290
68	Er	0.25716	0.25248	0.22260	0.21715	0.21566
69	Tm	0.24911	0.24436	0.21530	(0.2101)	0.2089
70	Yb	0.24147	0.23676	0.20676	0.20163*	0.20023
71	Lu	0.23405	0.22928	0.20212	0.19689	0.19584
72	Hf	0.22699	0.22218	0.19554	0.19031	0.18981
73	Ta	0.220290	0.215484	0.190076	0.18508 (D)	0.18393
74	W	0.213813	0.208992	0.184363	0.17950 (D)	0.17837
75	Re	0.207598	0.202778	0.178870	0.17415 (D)	0.17311

TABLE 3.1
Wavelengths of K-emission Series and Absorption Edges
The units are kX 1-00202. They may be taken as in Å to an accuracy of 1 in 25,000 only (see Section 2.1.2).

At. No.	Element symbol	A_{K_2}	A_{K_1}	$A_{K\beta_1}$	$A_{K\beta_2}$ (D signifies mean of doublet)	K absorption edge
1	H					
2	He					226.5
3	Li	240				
4	Be	113				
5	B	67				
6	C	44				43.5
7	N	31.60				30.2
8	O	23.71				23.2
9	F	18.31				
10	Ne	14.616		14.464		
11	Na	11.909		11.617		9.5117
12	Mg	9.8889		9.558		7.9511
13	Al	8.33916	8.33669	7.981		6.7446
14	Si	7.12773	7.12528	6.7681		5.7866
15	P		6.1549	5.8038		
16	S	5.37471	5.37196	5.03169		5.0182
17	Cl	4.73050	4.72760	4.4031		4.3969
18	Ar	4.19456	4.19162	3.8848*		3.8707
19	K	3.74462	3.74122	3.4538		3.43645
20	Ca	3.36159	3.35825	3.0896		3.07016
21	Sc	3.03452	3.03114	2.7795		2.7573
22	Ti	2.75207	2.74841	2.51381		2.49730
23	V	2.50729	2.50348	2.28434		2.26900
24	Cr	2.29351	2.28962	2.08480		2.07017
25	Mn	2.10568	2.10175	1.91015		1.89628
26	Fe	1.93991	1.93597	1.75653		1.74311
27	Co	1.79278	1.78892	1.62075		1.60611
28	Ni	1.66169	1.65784	1.50010	1.48861	1.48502
29	Cu	1.54433	1.54051	1.39217*	1.38102	1.38011
30	Zn	1.43694	1.43311	1.29522	1.28366	1.28311
31	Ga	1.34394	1.34013	1.20784	1.19595	1.19507
32	Ge	1.25797	1.25401	1.12890	1.11682	1.11600
33	As	1.17981	1.17581	1.05726	1.04498	1.04411
34	Se	1.10875	1.10471	0.99212	0.97986	0.97900
35	Br	1.04376	1.03969	0.93273	0.92064	0.91995
36	Kr	0.9841	0.9801	0.87845	0.86609	0.86547
37	Rb	0.92963	0.92551	0.82863	0.81641	0.81549
38	Sr	0.87938	0.87524	0.78285	0.77076	0.76969
39	Y	0.83000	0.82579	0.74068	0.72874	0.72762
40	Zr	0.78010	0.77588	0.70170	0.68989	0.68877
41	Nb	0.75040	0.74615	0.66572	0.65412	0.65291
42	Mo	0.713543	0.70926	0.632253	0.62099 (D)	0.61977
43	Tc	0.67927*	0.67493*	0.60141*	0.59018*	0.58911
44	Ru	0.64736	0.64304	0.57246	0.56164	0.56011
45	Rh	0.617610	0.613245	0.54559	0.53509 (D)	0.53320
46	Pd	0.589801	0.585415	0.52052	0.51021	0.50910
47	Ag	0.563775	0.559363	0.49701	0.48701	0.48584
48	Cd	0.53941	0.53498	0.475078	0.46531	0.46409
49	In	0.51652	0.51209	0.454514	0.444963	0.44388
50	Sn	0.49502	0.49056	0.435216	0.425900	0.42468
51	Sb	0.47479	0.470322	0.417060	0.407950	0.40663
52	Te	0.455751	0.451263	0.399972	0.391080	0.38972
53	I	0.437805	0.433293	0.383884	0.37547	0.37370
54	X	0.42043	0.41596	0.36846	0.35989	0.35849
55	Cs	0.404812	0.400268	0.354347	0.346084	0.34474
56	Ba	0.389646	0.385089	0.340789	0.332745	0.33137
57	La	0.375279	0.370709	0.327959	0.32024 (D)	0.31842
58	Ce	0.361665	0.357075	0.315792	0.30826 (D)	0.30647
59	Pr	0.348728	0.344122	0.304238	0.29690 (D)	0.29516
60	Nd	0.336487	0.331822	0.293274	0.28631	0.28451
61	Pm	0.3249	0.3207	0.28209	(0.2761)	(0.2743)
62	Sm	0.31365	0.30895	0.27305	0.26629	0.26462
63	Eu	0.30326	0.29850	0.26360	0.25697	0.25552
64	Gd	0.29320	0.28840	0.25445	0.24812	0.24680
65	Tb	0.28343	0.27876	0.24601	0.23960	0.23840
66	Dy	0.27430	0.26957	0.23758	0.23175	0.23046
67	Ho	0.26552	0.26081	(0.2302)	(0.2244)	0.22290
68	Er	0.25716	0.25248	0.22260	0.21715	0.21566
69	Tm	0.24911	0.24436	0.21530	(0.2101)	(0.2089)
70	Yb	0.24147	0.23676	0.20876	0.20363*	0.20223
71	Lu	0.23405	0.22928	0.20212	0.19689	0.19584
72	Hf	0.22699	0.22218	0.19554	0.19031	0.18981
73	Ta	0.220290	0.215484	0.190076	0.18508 (D)	0.18393
74	W	0.213813	0.208992	0.184363	0.17950 (D)	0.17837
75	Re	0.207598	0.202778	0.178870	0.17415 (D)	0.17311

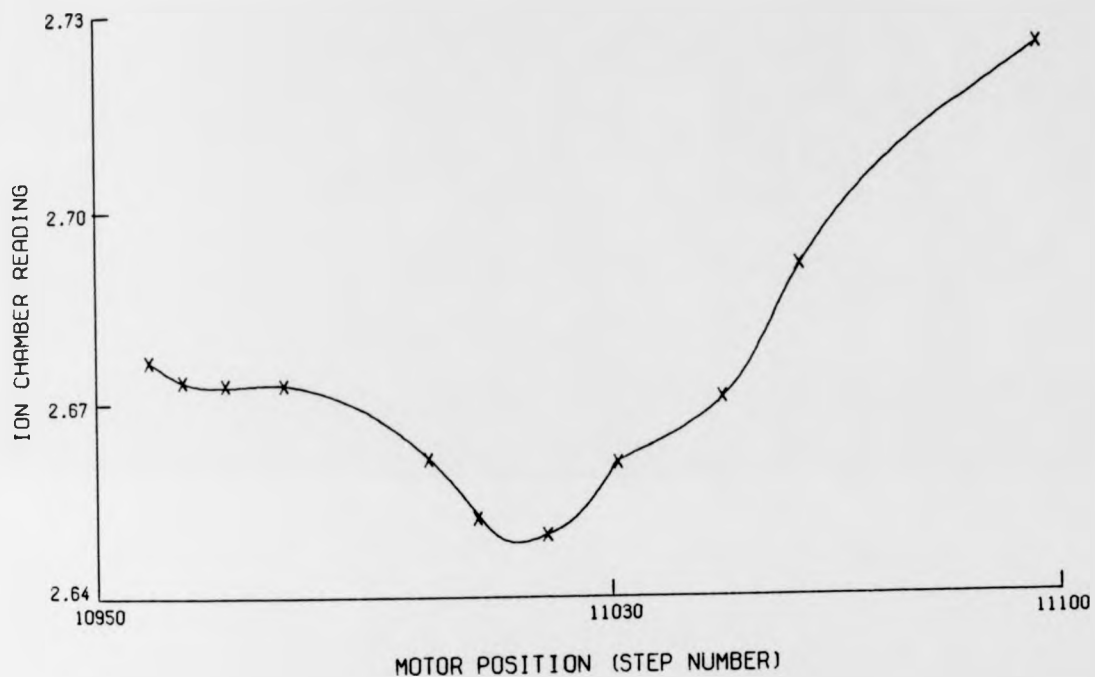


FIGURE 3.20 TRANSMISSION SPECTRUM OF COPPER K EDGE OBTAINED
WHILST SETTING THE TWO THETA ANGLE ON THE WORKSTATION

A detailed alignment procedure for aligning the top carriage to the beam is given by J.R. Helliwell and A.W. Thompson (1983). Once the beam is actually passing through the collimator the flux is maximised by a few cycles of the following procedure:

- (1) The ion chamber reading is maximised using vertical translations (key 23);
- (2) Maximise the reading using horizontal translations (key 21);
- (3) Adjust the tilt of the camera using the vertical back (key 19) and horizontal back (key 18) motors.
- (4) Remaximise vertical and horizontal translations.

3.5.7 Radiation Protection

It has been shown (Golde and Warren (1980)) that the major risk of exposure to radiation occurs at beam injection and that subsequent synchrotron radiation from stable stored electron beams (up to energies of 3.5 GeV) are easily shielded by 1/8th inch sheet metal. For this reason no access at all is permitted into the experimental areas during beam injection and subsequent energy ramping of the stored beam. Once the required energy and currents have been established access to the experimental areas is permitted. The protein crystallography workstation is enclosed in its own local radiation enclosure (hutch). No access to the inside of the hutch is allowed whilst X-rays are entering it. The X-rays can be shut off in one of two

ways, firstly the main port shutter which is operated by the control room and affects the EXAFS, protein crystallography and topography workstations. Secondly, a pair of local shutters solely for the protein crystallography workstation are under the direct control of the experimenter.

The local shutters are interlocked to the hutch doors and beam cannot be allowed into the beamline until the hutch has been searched (and found to be empty of people!). The user is forced to look around the hutch whilst walking between push button switches that form the search points (see figure 3.15(i)). The search is performed as follows:

- (1) A key is taken from the switch in the mask control module (situated outside the hutch in the p.x. control rack).
- (2) The 'START SEARCH' button is pressed on the search control box to the left of the main entrance to the hutch. Once this button has been pressed there is approximately 40 seconds in which to complete the search.
- (3) The key obtained in step (1) is then inserted in a lock at the first search position (see figure 3.14(a)). This enables a button which can then be pressed.
- (4) The experimenter proceeds to the second search position where he/she presses another button.
- (5) The button at the third position (by the exit to

the hutch) is then pressed.

- (6) The hutch door is gently closed and locked (with the searcher outside!).
- (7) The key must then be returned to the mask control module and turned.
- (8) An audible alarm then sounds for approximately 20 seconds, after which time the port can be opened and closed at will by the experimenter via the 'open' and 'close' buttons in the control rack.

A similar but larger scale search is made of all experimental areas prior to beam injection.

3.6 Conclusion

The perfect crystal monochromator/platinum coated quartz mirror focussing instrument, used for protein crystallographic experiments on the SRS Daresbury, U.K., has been discussed in detail. It has been shown that this is a highly versatile and precise instrument that has been used to quickly collect experimental data that was unattainable from 'conventional' X-ray sources (see section 7.2).

The relevant theory has been cited that allows calculation of beam compression, Guinier position and spectral resolution for this instrument. In addition the theory required to optimise the beam geometry in terms of

flux and the size of Bragg reflections have been derived from first principles. The effect of the depth of focus is also discussed (see section 3.4.5). A formula for calculating the position of the experimental arm for a required wavelength (2θ) setting has been derived, and a simple procedure for aligning the camera to the synchrotron beam was also given in this chapter.

CHAPTER 4

DATA ACQUISITION AND REDUCTION

4.1 Data Collection

Bacterial 6PGDH crystallises into small, (0.15mm maximum crystal size), trigonal crystals in the space group $P3_221$ (see (figure 4.1)). The molecular dimer is the asymmetric unit. The unit cell dimensions are:

$$\underline{a} = 123.1 \text{ \AA}$$

$$\underline{b} = 123.1 \text{ \AA}$$

$$\underline{c} = 147.7 \text{ \AA}$$

$$\alpha = 90^\circ$$

$$\beta = 90^\circ$$

$$\gamma = 120^\circ$$

The Laue group ($\bar{3}m$) is shown in figure 4.2 along with the unit cell directions for a \underline{c}^* mount with \underline{a} antiparallel to the X-ray beam. This was the mount used for all our crystals. Due to the trigonal symmetry we required to collect sixty degrees of data in order to obtain four equivalents of data. The angular ranges of data collected on the SRS are shown for all our crystals in table 4.1

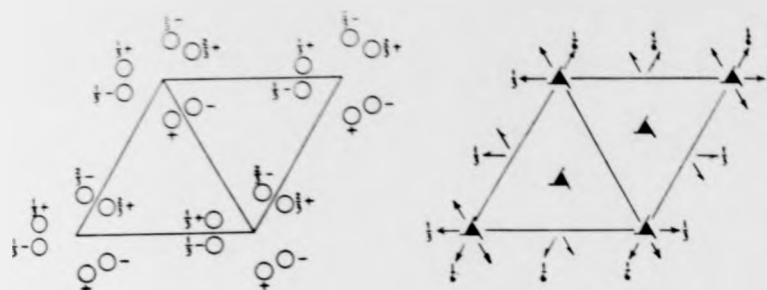
Unfortunately the data from crystals DNB1, DNB2 and DNB3 were collected on a batch of Kodak Noscreen which

Trigonal 32

$P3_221$

No. 154

$P3_221$
 D_3^6



Origin at 3_221 [2-axis normal to (1120)]

Number of positions,
Wyckoff notation,
and point symmetry

Co-ordinates of equivalent positions

Conditions limiting
possible reflections

6 c 1 $x, y, z; \bar{y}, x - y, \frac{1}{2} + z; y - x, \bar{x}, \frac{1}{2} - z;$
 $y, x, \bar{z}; \bar{x}, y - x, \frac{1}{2} - z; x - y, \bar{y}, \frac{1}{2} - z.$

General:

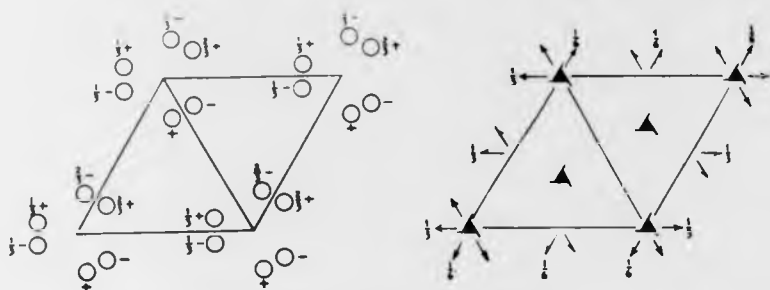
hkl: No conditions
000: $l = 3n$

Special: as above only

3 b 2 $x, 0, \frac{1}{2}; 0, x, \frac{1}{2}; \bar{x}, \bar{x}, \frac{1}{2}.$

3 a 2 $x, 0, \frac{1}{2}; 0, x, \frac{1}{2}; \bar{x}, \bar{x}, 0.$

Figure 4.1 The Space Group $P3_221$

Origin at $3_2 21$ [2-axis normal to (1120)]Number of positions,
Wyckoff notation,
and point symmetry

Co-ordinates of equivalent positions

Conditions limiting
possible reflections

6	c	1	$x, y, z; \bar{y}, x - y, \frac{2}{3} + z; y - x, \bar{x}, \frac{1}{3} + z;$ $y, x, \bar{z}; \bar{x}, y - x, \frac{1}{3} - z; x - y, \bar{y}, \frac{1}{3} - z.$
---	---	---	--

General:

 hkl : No conditions
 $000l$: $l = 3n$

Special: as above only

3	b	2	$x, 0, \frac{1}{2}; 0, x, \frac{1}{2}; \bar{x}, \bar{x}, \frac{1}{2}.$
---	---	---	--

3	a	2	$x, 0, \frac{1}{2}; 0, x, \frac{1}{2}; \bar{x}, \bar{x}, 0.$
---	---	---	--

Figure 4.1 The Space Group $P3_2 21$

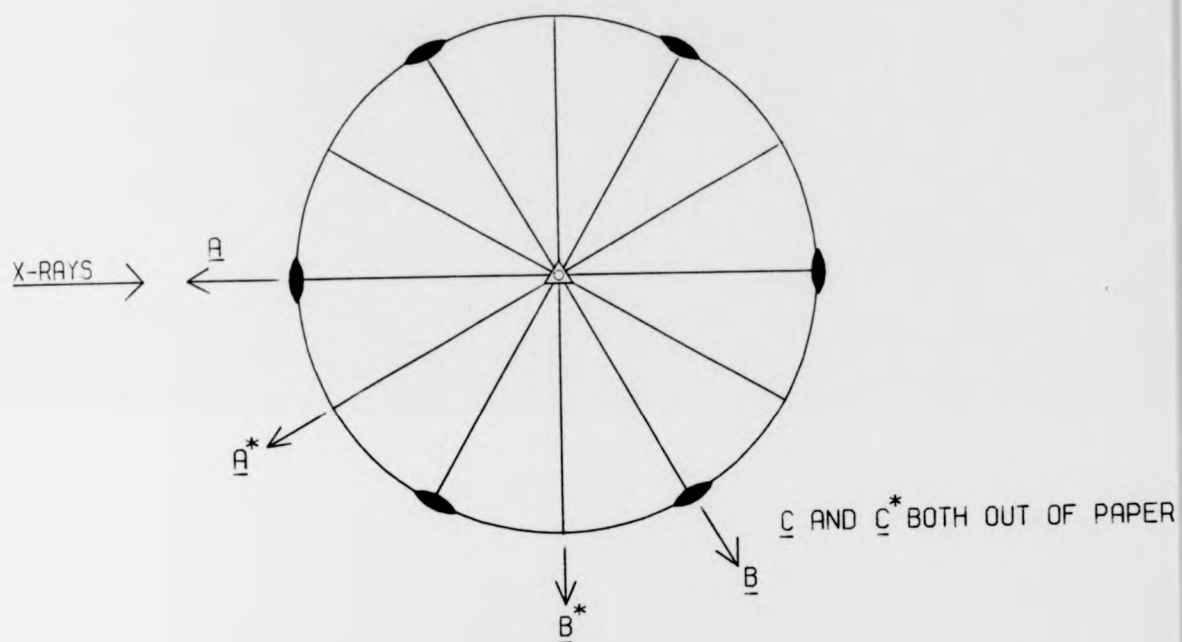


FIGURE 4.2 LAUE GROUP $\bar{3}$, SHOWING UNIT CELL DIRECTIONS FOR A
C* MOUNT

Table 4.1 Angular ranges and notes for the DNB range of crystals.

CRYSTAL	ANGULAR RANGE (DEGREES)	NOTES
DNB1	-1.5 to 10.5	Bad batch of Kodak film used. Data had $R_{\text{sym}} = 51.2\%$ to 3.33% . Data not used.
DNB2	9.0 to 24.0	
DNB3	22.5 to 40.5	
DNB4	39 to 57	CEA film pre-mirror data
DNB5	55.5 to 60	CEA film pre-mirror data
DNB6	0 to 15	Not used.
DNB7	28 to 40	CEA film post-mirror data
DNB8	10 to 29.5	CEA film post-mirror data

appears to have been faulty, giving an uneven darkening of the film for equivalent reflections. The processed data had an R_{sym} value (see section 4.2.7) of 51.2% at a resolution of 3.33 Å. For this reason the data was rejected and only data from crystals DNB4, DNB5, DNB7 and DNB8 were used in the final dataset. DNB6 was not processed because data from the other four crystals were sufficient for a full dataset. There was also a great deal of competition for scanning and computing time when the data was being processed. Figure 4.3 shows the data that were collected for this dataset.

4.2 Data Processing

The data reduction and processing was done using the MOSCO suite of programs. (A.J. Wonacott (1980)). The programs had been modified to allow them to run off-line, from a magnetic tape of digitised data, rather than on-line to a scanning microdensitometer (Machin and Greenhough (1983)). Personal experience with earlier datasets proved that this suite of programs run on a SYSTIME 8750 computer at Daresbury Laboratory was much faster and easier to use than the OSCAR suite of programs (Wilson and Yeates (1979)) running on the mainframe IBM/370 computer at Daresbury.

4.2.1 Refinement of crystal orientation and unit cell parameters

The program IDXREF was used for refining the crystal orientation and unit cell parameters. Still photographs were scanned on a Joyce-Lobell Scandig 3 scanning microdensitometer. The position of 75 spots per film relative to the fiducial marks were recorded using the program STILLs. The position of these spots were input into IDXREF and their partiality assumed to be 0.5

IDXREF indexes the spots using the method reported by Schwager, Bartels and Jones (1975). Then, using approximate input values of unit cell, orientation missetting angles (measured from setting stills) and the fiducial coordinates relative to the centre of the film, the program refines the misorientation angles θ_x , θ_y , and θ_z also the cell parameters. The refinement is a least squares refinement minimising the function:

$$\sum_i (d_i)^2 = \sum_i \left(\frac{R_o - R_c}{d^*} \right)^2$$

where the summation i is over all reflections,

R_c is the calculated distance from the centre of the Ewald sphere to the reciprocal lattice point,

R_o is the observed distance from the centre of the Ewald sphere to the reciprocal lattice point.

For the observed distance R_o , the lattice point is imagined to have a finite diameter of magnitude $d^* \Delta \cos \theta$, where Δ is the total reflecting range of the crystal for reflections on the equatorial plane; it is the sum of the angular dispersion of the X-ray beam γ , the crystal mosaicity η and the wavelength dispersion $\frac{\delta \lambda}{\lambda} \tan \theta$, d^* is the length of the reciprocal lattice vector ($= 2 \sin \theta$). Each residual $|d_i|$ is checked to see if it is below a user specified limit, if it is not that particular reflection is omitted from the refinement.

The refinement continues for three cycles or until the shifts in all the parameters being refined are less than 10% of the standard deviation of that parameter. Upon convergence or three cycles the final values of the refined parameters are listed and the orientation matrix [A] is output. The refined orientation matrix and missetting angles are then used to predict the diffraction pattern (see section 4.2.2) for the first two and last two packs of the crystal data. The post refinement program POSTCHK (see section 4.2.4) is then used to obtain measured partialities for these films. These measured partialities can then be used to further refine the orientation matrix and missetting angles if necessary by inputting them to IDXREF. The program outputs a rms value for the residual angle (in degrees) between the observed and calculated position of the reciprocal lattice point subtended at the lattice origin. It also outputs the final value for the residual $\sum_1 (d_i)^2$

given above. Table 4.2 shows the refined values of the missetting angles, reciprocal cell constants and residuals obtained for bacterial 6PGDH using IDXREF.

4.2.2 Prediction of reflections and their partiality

In order to utilise the rotation method of data collection efficiently one must be able to predict the reflections which will appear on films of any particular ϕ (spindle) range. It would be very inefficient and time consuming to test every reciprocal lattice point to see whether it falls into a particular θ range for each film: a 100 Å unit cell to a resolution of 2.5 Å would require testing nearly 250,000 reflections to extract about 2,500 reflection recorded on one film (Wonacott (1977)).

The relationship between the reciprocal lattice coordinates (xyz) and the Miller indices (hkl) is:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = [\phi] [C] [A] \begin{bmatrix} h \\ k \\ l \end{bmatrix}$$

where [C] is the combined matrix for misorientation angles ϕ_x, ϕ_y, ϕ_z .

Therefore for the two extreme positions of ϕ at the start (ϕ_S) and the end (ϕ_E) of a rotation are given by:

Table 4.2 Summary of IDREF results for the bacterial 6PGDH data.

CRYSTAL	DNB1	DNB2	DNB3	DNB4 [†]	DNB5 [†]	DNB7 [†]	DNB8 [†]
$\phi x(^{\circ})$	0.09196	-0.16404	-0.03011	-3.40926	0.65103	0.77892	-1.82846
$\phi y(^{\circ})$	-0.15583	-0.02690	0.13146	0.24391	-0.30936	-0.0876	-2.71050
$\phi z(^{\circ})$	-29.63	-29.55576	-29.6727	-30.239	-29.78516	-29.9797	-29.6307
$a^{\circ}(\text{\AA}^{-1})$	0.013948	0.013968	0.013770	0.013980	0.013947	0.014040	0.013969
$b^{\circ}(\text{\AA}^{-1})$	0.013948	0.013968	0.013770	0.013980	0.013947	0.014040	0.013969
$c^{\circ}(\text{\AA}^{-1})$	0.01011	0.010059	0.01081	0.01008	0.010076	0.010170	0.010217
$[(R_0 - R_C)/d]^2$	6.94×10^{-5}	9.06×10^{-5}	8.57×10^{-5}	9.11×10^{-5}	8.20×10^{-5}	2.93×10^{-4}	4.66×10^{-4}
rms resid (degree)	0.044	0.057	0.122	0.129	0.057	0.080	0.126

[†] denotes used in final dataset

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_S = [\phi]_S [C] [A] \begin{bmatrix} h \\ k \\ l \end{bmatrix}$$

and

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}_E = |\phi|_E |C| |A| \begin{bmatrix} h \\ k \\ l \end{bmatrix}$$

The distance of the point $\begin{bmatrix} x \\ y \\ z \end{bmatrix}_S$ from the centre of the Ewald sphere is given by:

$$r_S^2 = (1-x_S)^2 + y_S^2 + z_S^2$$

and similarly for r_E .

Therefore the condition for the point to have passed through the Ewald sphere is given by:

$$r_S > 1 ; \quad r_E < 1$$

As was mentioned before (see section 4.2.1), the reciprocal lattice point is expanded to account for the reflecting range of the crystal and can be assumed to be a spherical volume of radius ϵ given by:

$$\epsilon = \frac{\Delta d^* \cos \theta}{2}$$

The test for a reciprocal lattice volume to have passed completely through the sphere then becomes:

$$r_S > 1 + \epsilon \quad ; \quad r_E < 1 - \epsilon$$

See figure 4.4.

The program used for the prediction of reflections and partialities in the MOSCO suite of programs is called GENVEE. A scan-step-jump algorithm is used to scan through the reciprocal lattice in an area close to the Ewald sphere. The reciprocal lattice points predicted by this algorithm then undergo a series of tests to determine their partiality. A tree diagram of these tests is shown in figure 4.5. Further tests decide whether reflections are overlapping, outside the physical limits of the film, or beyond a user specified resolution limit. The program outputs a 'generate file' containing the following information for each reflection.

$$h \quad k \quad l \quad m \quad r \quad X_f \quad Y_f$$

where hkl = Miller indices,

m = index of partiality, ($m = 0$ fully recorded;

$m = 1$ to 8 partially recorded).

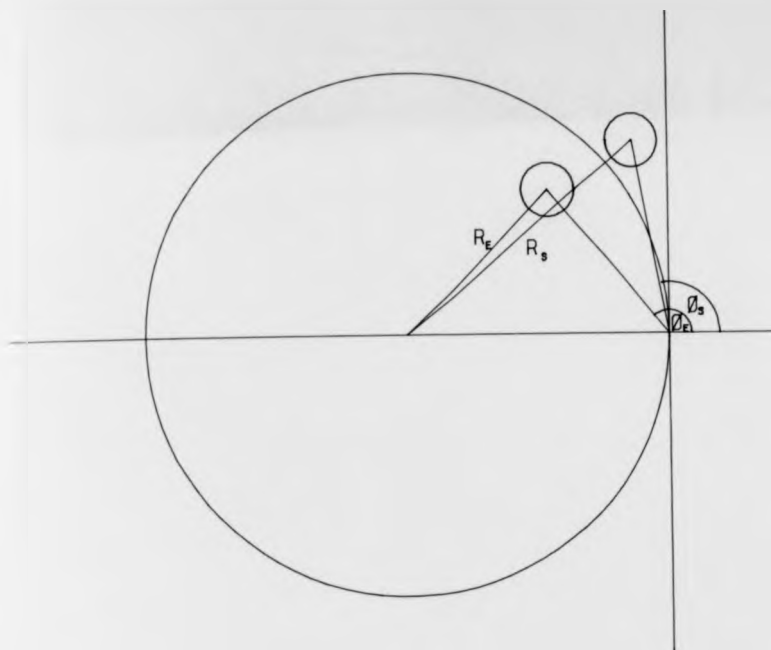


FIGURE 4.4 PROJECTION DOWN THE ROTATION AXIS SHOWING THE MOVEMENT OF RECIPROCAL LATTICE VOLUMES THROUGH THE SPHERE OF REFLECTION DURING A ROTATION $\theta_E - \theta_S$

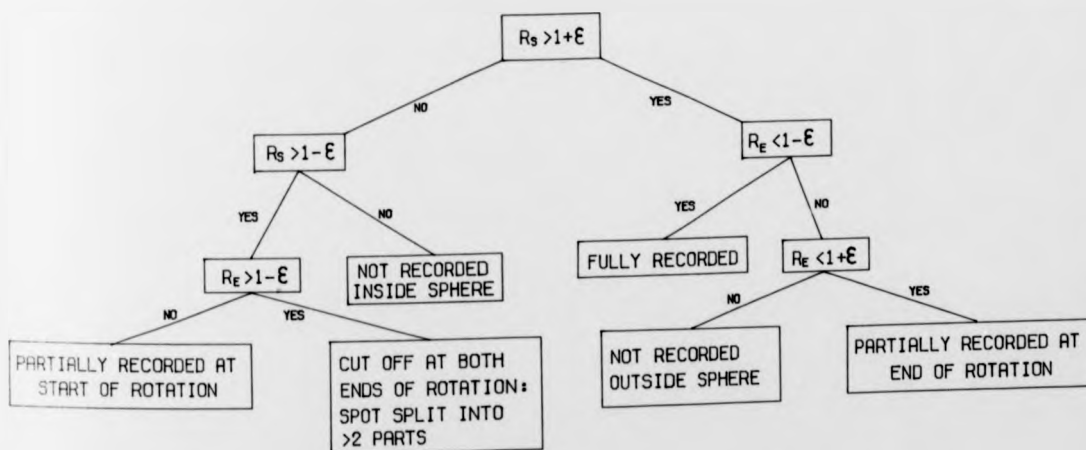


FIGURE 4.5 TREE DIAGRAM OF TESTS USED TO ACCERTAIN THE REFLECTING CONDITION OF A RECIPROCAL LATTICE VOLUME

r = measurability flag, ($r = 0$ measurable; $= 1, 2, 3$ failed a test).

x_f, y_f = film coordinates of reflection.

Space is reserved at the end of each reflection record for the insertion of the integrated intensity and standard deviations for up to three films per pack by a later program in the suite (see section 4.2.3). An image of the film predicted is output to a graphics terminal for visual comparison with the data film. The program will predict as many films as are required by the user up to a maximum of 32,768 reflections.

4.2.3 Recording of integrated intensities and background measurements

The films containing the diffraction data were all scanned on a rotating drum scanning microdensitometer (Joyce-Lobel Scandig 3) at a raster of fifty microns. The digitised image of the films were written to magnetic tape and a modified off-line version of the program MOSFLM (Machin and Greenhough (1983)) was used to extract integrated intensities and background measurements. A flow chart of the program is shown in figure 4.6. The handswitch options (hsw3 etc.) shown on the flow chart are replaced by responses to questions from the terminal in the off-line version of the program.

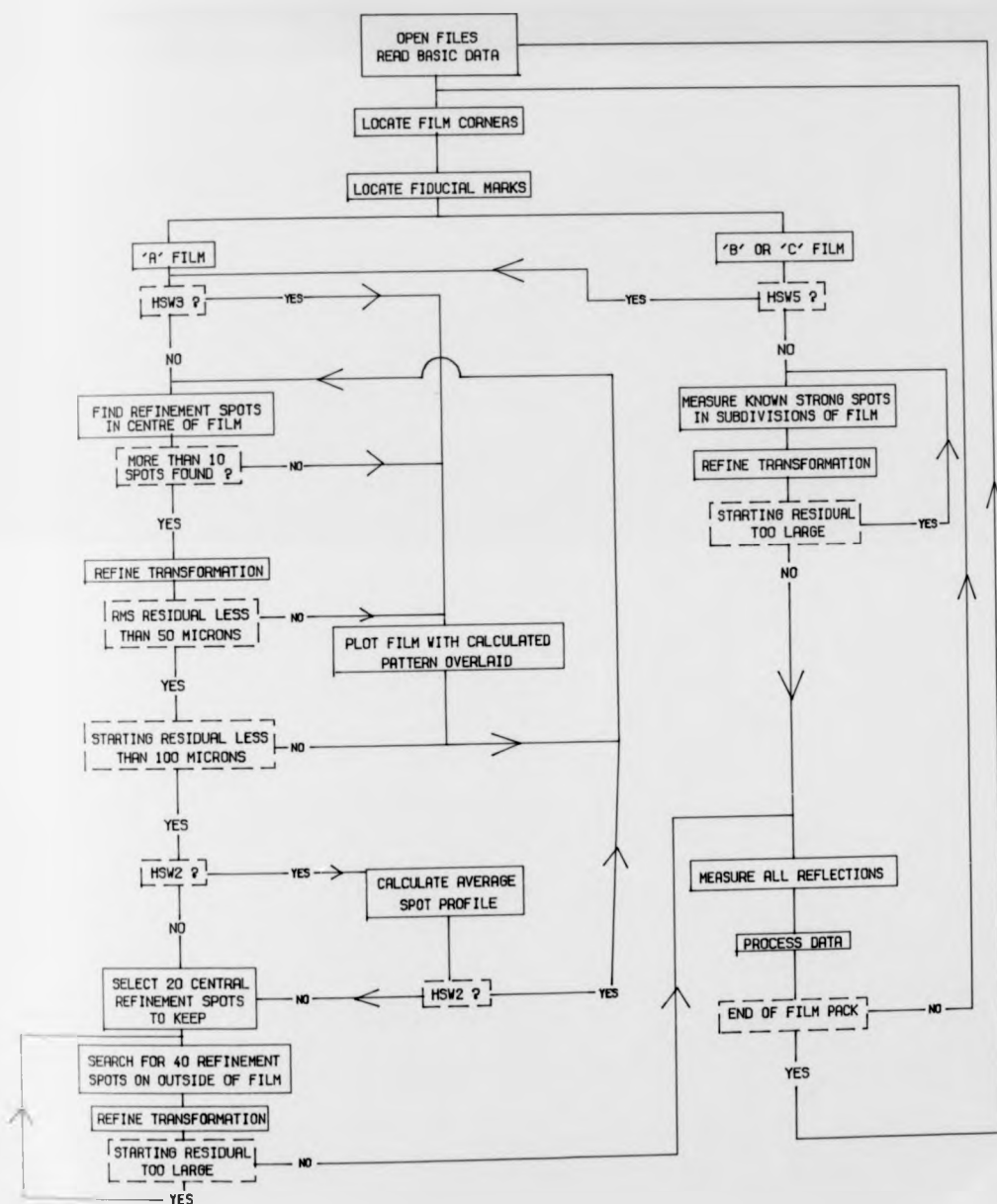


FIGURE 4.6 FLOW DIAGRAM OF 'MOSFLM'

The first step in the program is to calculate transformations between the scanners frame of reference (the digital image of the film being recorded relative to the scanners frame) and the film coordinate system. The position of the fiducial marks are used to do this. The relationship between the two frames of reference are defined by X_{CEN} , Y_{CEN} , and ω as shown in figure 4.7. The actual transformation is given by:

$$X_d = F(X_{CEN} + K(X_f \cos \omega - Y_f \sin \omega))$$

$$Y_d = F(Y_{CEN} + K Y_{SCA} (X_f \sin \omega + Y_f \cos \omega))$$

where X , Y , X_{CEN} , Y_{CEN} and are defined in figure 4.7,

F = factor to convert 10 micron units into scanner units,

K = overall scale factor (≈ 1 as crystal to film distance is implicit in the film coords),

Y_{SCA} defines relative scales along Y_f and X_f

These transformations are refined to allow for film distortions by means of replacing K in the above expression by K' as given by:

$$K' = K + \text{TILT } X_f + \text{TWIST } Y_f + \text{BULGE } R_f$$

These film distortion parameters allow corrections to be made either due to the film plane not being normal to the

The first step in the program is to calculate transformations between the scanners frame of reference (the digital image of the film being recorded relative to the scanners frame) and the film coordinate system. The position of the fiducial marks are used to do this. The relationship between the two frames of reference are defined by X_{CEN} , Y_{CEN} and ω as shown in figure 4.7. The actual transformation is given by:

$$X_d = F(X_{CEN} + K(X_f \cos \omega - Y_f \sin \omega))$$

$$Y_d = F(Y_{CEN} + K Y_{SCA} (X_f \sin \omega + Y_f \cos \omega))$$

where X , Y , X_{CEN} , Y_{CEN} and are defined in figure 4.7,

F = factor to convert 10 micron units into scanner units,

K = overall scale factor (≈ 1 as crystal to film distance is implicit in the film coords),

Y_{SCA} defines relative scales along Y_f and X_f

These transformations are refined to allow for film distortions by means of replacing K in the above expression by K' as given by:

$$K' = K + \text{TILT } X_f + \text{TWIST } Y_f + \text{BULGE } R_f$$

These film distortion parameters allow corrections to be made either due to the film plane not being normal to the

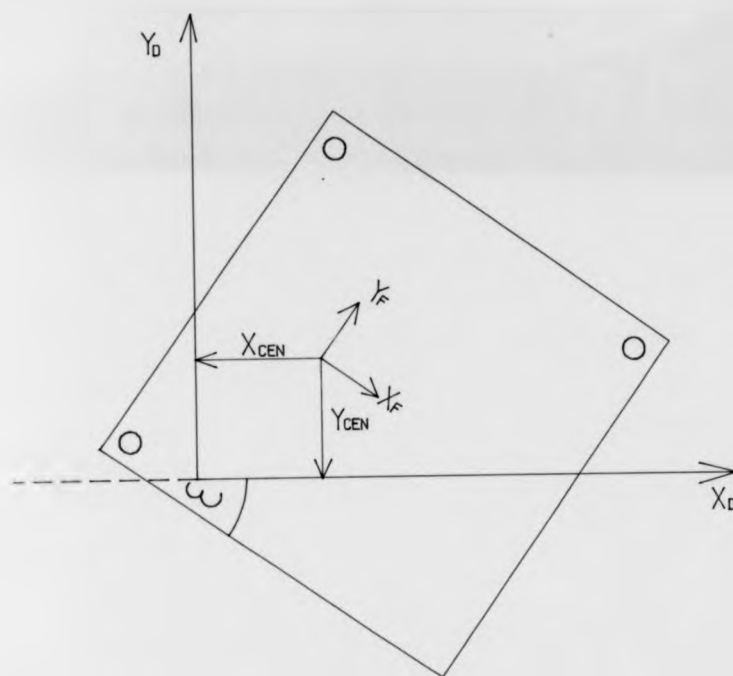


FIGURE 4.7 RELATIONSHIP BETWEEN THE SCANNER (X_D, Y_D) AND FILM (X_F, Y_F) FRAMES OF REFERENCE

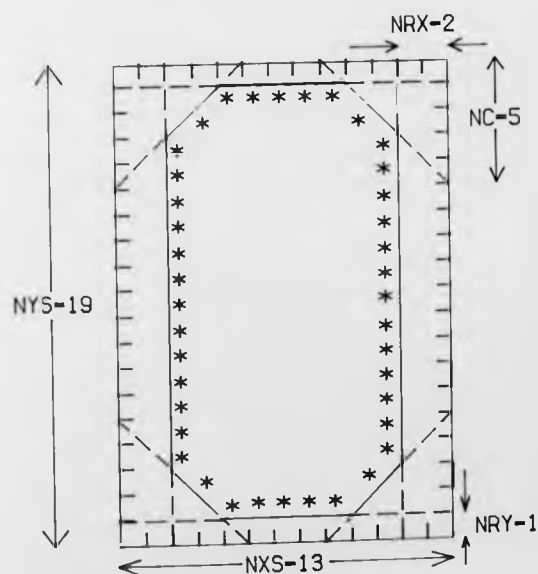


FIGURE 4.8 THE INTEGRATION BOX AS USED IN 'MOSFLM'

X-ray beam or the film not being flat in the cassette. The bulge parameter allows for the film bulging in the cassette or not being correctly secured to the scanner drum.

Using these transformations the predicted pattern from the coordinates output from GENVEE (see section 4.2.2) are shown on a graphics terminal overlayed with an image of the observed diffraction pattern. If necessary a cursor can be used to locate equivalent positions on the two images. This allows further refinement of the transformations. When one is happy with the coincidence of the predicted and observed patterns visually, the program selects up to sixty spots from a central five centimeter square region of the film. The transformations are again refined using the positions of these spots. The refinement is repeated until the rms residual is less than fifty microns.

Once the refinement is satisfactory the program optionally calculates an average spot profile for the spots located in the central area of the film. The profile is displayed on the terminal and any alterations to the measurement box can be made if it becomes apparent that the one currently being used is not optimal. The measurement box is defined by six parameters (see figure 4.8) which allow the user to vary the way in which the spot peak intensity and background are measured. The box used for bacterial 6PGDH had the dimensions:

NXS = 13 raster units,

NYS = 19 raster units,

NC = 5 raster units,

NRX = 2 raster units,

NRY = 1 raster unit.

This gave 106 raster points in the peak, 106 raster points in the background with 40 raster points not recorded. The 40 unrecorded points form the boundary between the peak and background areas and help to avoid the problem of peaks 'smearing' into the background.

A final refinement of the transformations is done using up to an additional 40 reflections from the high angle data.

The integrated intensities are then measured for all the reflections predicted in the 'generate' file. The measurement box expands automatically to account for the oblique incidence at higher angles. The standard deviations of both the peak and background measurements are calculated along with the ratio of peak/background and shifts in the centre of gravity of the peaks from their predicted positions. Any reflections which have unduly large shifts or background ratios are flagged and may be inspected and/or rejected from the dataset.

The program writes the integrated intensities and standard deviations back to the 'generate' file. Finally a

vector plot of the shifts of the peak centres of gravity from the expected positions for all the measured reflections is shown on a graphics terminal. This shows up any systematic errors caused by wrong positioning of the measurement box or incorrect crystal to film distance refinement.

4.2.4 Post-refinement

The program POSTCHK was used on the first two and last two packs of each crystal. The program reads the intensities of those reflections which are partial at a particular ϕ value (at the interface of two adjacent films) and determines the observed degree of partiality, from the integrated intensities recorded in the 'generate' file. Statistics are collected on the reflections and if a significant amount are not acceptably close to the predicted partialities from GENVEE, then the indices and partialities are input to IDXREF for a post-refinement of the orientation parameters (see section 4.2.1).

The program also displays a graphical representation of the spots showing whether the measured partialities are too high, too low or acceptable. By processing the first two and last two packs of a crystal and running POSTCHK one can easily tell if the orientation matrix is acceptable for the whole ϕ range or if there has been any crystal slippage.

vector plot of the shifts of the peak centres of gravity from the expected positions for all the measured reflections is shown on a graphics terminal. This shows up any systematic errors caused by wrong positioning of the measurement box or incorrect crystal to film distance refinement.

4.2.4 Post-refinement

The program POSTCHK was used on the first two and last two packs of each crystal. The program reads the intensities of those reflections which are partial at a particular ϕ value (at the interface of two adjacent films) and determines the observed degree of partiality, from the integrated intensities recorded in the 'generate' file. Statistics are collected on the reflections and if a significant amount are not acceptably close to the predicted partialities from GENVEE, then the indices and partialities are input to IDXREF for a post-refinement of the orientation parameters (see section 4.2.1).

The program also displays a graphical representation of the spots showing whether the measured partialities are too high, too low or acceptable. By processing the first two and last two packs of a crystal and running POSTCHK one can easily tell if the orientation matrix is acceptable for the whole ϕ range or if there has been any crystal slippage.

4.2.5 Correction of integrated intensities for geometrical factors

The following corrections are made to the integrated intensities due to geometry of the rotation camera:

- (a) Oblique incidence factor,
- (b) Lorentz factor,
- (c) Polarisation factor.

These corrections were applied using the program ROTCOR. The program inputs data from one or more 'generate' file(s) and outputs a separate file ready for film to film scaling. The output records are of the form:

$$h \ k \ l \ m \ I_A \text{ ISD}_A \ I_B \text{ ISD}_B \ I_C \text{ ISD}_C \ d^2 L_p^{-1}$$

where I_A, I_B, I_C are the integrated intensities,
 $\text{ISD}_A, \text{ISD}_B, \text{ISD}_C$ are their standard deviations,
 L_p^{-1} is the Lorentz-polarisation factor.

(a) Oblique incidence at the film is corrected using the expression:

$$S_n(\beta) = A(\beta) B(\beta) C(\beta)^{n-1}$$

where $S(\beta)$ is the factor by which intensities are multiplied on the n th film in the pack.

$A = 2\theta$ for flat films.

$$A(\beta) = a(0)/a(\beta)$$

and

$$a(\beta) = (1 - \exp(-k_E \sec \beta)) / (1 + \exp(-k_E - k_B) \sec \beta)$$

$$B(\beta) = \exp(-k_P) / \exp(-k_P \sec \beta)$$

$$C(\beta) = \exp(-k_f) / \exp(-k_f \sec \beta)$$

k_E , k_B , k_P are absorption coefficients for the film emulsion, film base and front paper respectively.

$k_f = (2k_E + k_B)$ is the total film absorption.

(b) The Lorentz factor is a factor which allows for the differing times spent in the reflecting position by different reciprocal lattice points. It is given by L in the expression.

$$I \propto \frac{I_{OBS}}{L}$$

$$\text{where } L^{-1} = (\sin^2 2\theta + \zeta^2)^{\frac{1}{2}}$$

ζ is the cylindrical polar coordinate corresponding to the z_0 axis.

(c) The Polarization correction for a synchrotron plus focussing perfect crystal monochromator system such as ours is discussed by Kahn et al (1982) and is given by:

$$P = P_0 - P'$$

$$\text{where } P_0 = \frac{1}{2} (1 + \cos^2 2\theta)$$

$$\text{and } P' = \frac{r'}{2} \cos 2\rho \sin 2\theta$$

r' is a machine and monochromator dependent parameter

(for full expression see Kahn et al (1982)).

ρ is the azimuthal angle in the film.

At the SRS for the 7.2 instrument τ' pre-mirror and post-mirror is 0.96 and 0.85 respectively (Duke and Helliwell unpublished results).

4.2.6 Film to film scaling

The corrected data from ROTCOR was input into an interfilm scaling program called PASCAL. The program determined scale factors between the films from the same pack. Using these scale factors, corrections were made for strong reflections where the front film(s) were saturated. The program also carried out analysis of agreement between intensities on different films of the same pack and between symmetry related reflections. The agreement between symmetry related reflections is expressed in terms of an R_{sca} value given by:

$$R_{sca} = \frac{\sum_{hkl} \sum_i |\bar{I} - I_i|}{\sum_{hkl} N\bar{I}}$$

where \bar{I} is the mean intensity of N symmetry related reflections.

A summary of the scale factors and R_{sca} values is shown in Table 4.3. To avoid errors in the calculation of scale

Table 4.1 Interfilm scaling results from PASCAL

CRYSTAL PACK	DNB1 A/B Scales A/C R	DNB2 A/B Scales A/C R	DNB3 A/B Scales A/C R	DNB4* A/B Scales A/C R	DNB5* A/B Scales A/C R	DNB7* A/B Scales A/C R	DNB8* A/B Scales A/C R
1	3.07 - 17.3	2.82 - 18.1	3.13 - 24.4	2.6 - 6.4	2.5 - 6.5	2.18 4.82 8.0	2.43 5.78 8.4
2	3.07 - 18.2	3.15 - 31.5	2.38 - 22.7	2.64 - 8.9	2.67 - 9.7	2.27 5.24 7.9	2.37 5.09 9.3
3	3.08 - 19.0	3.26 - 36.7	2.79 - 21.7	2.72 - 7.1	2.71 - 8.4	2.34 5.53 9.3	2.45 5.4 7.9
4	3.02 - 17.5	3.23 - 27.3	3.08 - 21.2	2.66 - 6.5		2.43 5.62 8.1	2.42 5.52 8.9
5	2.94 - 18.0	3.08 - 26.0	3.19 - 26.0	2.57 - 5.8		1.95 5.4 8.7	2.46 5.61 9.3
6	2.98 - 19.6	3.33 - 29.4	3.27 - 30.4	2.59 - 6.8		2.42 5.77 7.7	2.14 4.77 7.8
7	2.95 - 18.6	3.34 - 30.9	3.04 - 24.0	2.65 - 7.6		1.89 4.14 8.2	2.34 5.41 7.5
8	2.98 - 17.0	4.77 - 32.8	3.95 - 26.9	2.65 - 8.4		2.33 5.21 6.9	2.21 5.18 8.1
9		3.83 - 28.9	3.11 - 26.6	2.65 - 6.8			2.17 5.11 8.1
10			3.06 - 20.4	2.59 - 3.7 ⁶			2.32 5.25 8.9
11			2.94 - 20.3	2.65 - 3.0 ⁶			2.07 4.12 8.3
12			3.12 - 21.5	2.65 - 2.6 ⁶			2.19 4.89 8.6

* denotes those crystals used in the final data set

⁶ denotes only data to 6Å scaled

factors due to weak reflections, a low intensity cut off was used. Reflections with an integrated intensity of lower than 300 were rejected from the scaling. The intensity scale was 0-255 per raster point for two optical densities.

4.2.7 Crystal to crystal scaling

The crystal to crystal scaling was done using the programs ROTAVATA and AGROVATA. ROTAVATA calculates scale factors and temperature factors for each batch (filmpack) of data, using the method of Fox and Holmes (1966). The scale factor for the i th batch given by:

$$K(i) = C(i) \exp \left(-2B(i) \frac{\sin^2 \theta}{\lambda^2} \right)$$

where $C(i)$ is a constant scale factor for the filmpack,

$B(i)$ is a temperature factor which accounts for the intensity variation with $\frac{\sin^2 \theta}{\lambda^2}$ due to thermal vibrations.

The individual scale and temperature factors are calculated with respect to the first pack in the dataset which is arbitrarily set to have a scale factor C of 1.0 and a temperature factor B of 0.0. Graphs showing the variation of B and C with exposure to the X-ray beam are shown in figure 4.9. The graphs have all been normalised so that the first batch of each crystal has a scale of 1.0 and a temperature factor of 0.0

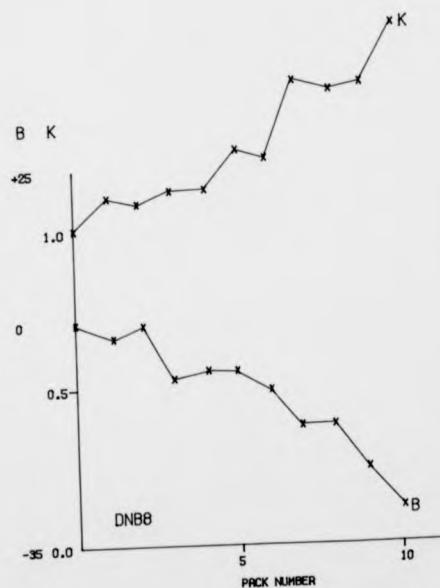
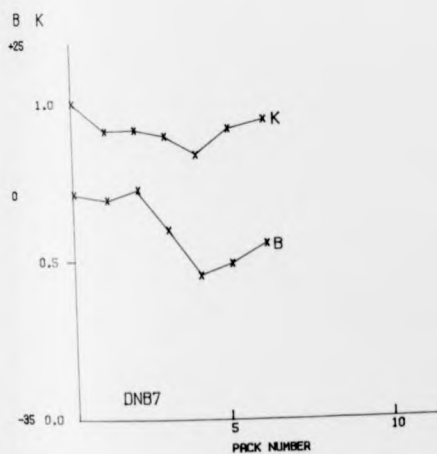
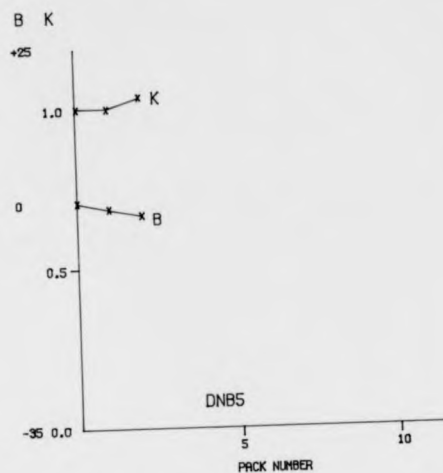
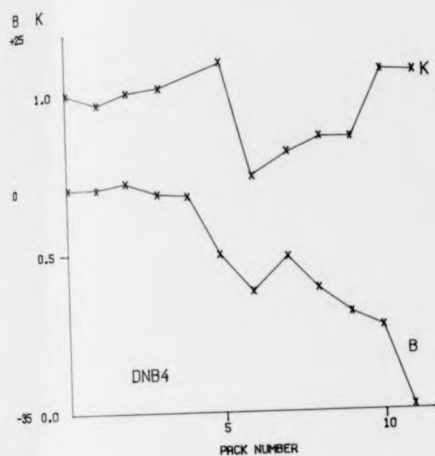


FIGURE 4.9 SCALE AND TEMPERATURE FACTORS AS CALCULATED
USING THE PROGRAM ROTAVATA

The program AGROVATA applies the scale factors calculated by ROTAVATA to the data. It also adds together partially recorded reflections, monitoring and rejecting bad agreements between repeated measurements and/or symmetry equivalents. AGROVATA also undertakes various analyses of the scaling including:

- (1) Analysis of scaling on a batch basis and overall,
- (2) Analysis of scaling as a function of $4\sin^2\theta/\lambda^2$
- (3) Analysis of scaling as a function of Intensity,
- (4) Analysis of partial bias,
- (5) Analysis of standard deviations.

Graphs showing the variation of R-factor with intensity and $4\sin^2\theta/\lambda^2$ are shown in figures 4.10 and 4.11. The R-factor used was:

$$R_{\text{sym}} = \frac{\sum_{hkl} \sum_{i=1}^N |\bar{I} - I_i|}{\sum_{hkl} \sum_{i=1}^N I_i}$$

where \bar{I} is the mean intensity of the N reflections with intensities I_i and common indices hkl .

4.2.8 Radiation Damage

The cause of radiation damage by X-rays to biological

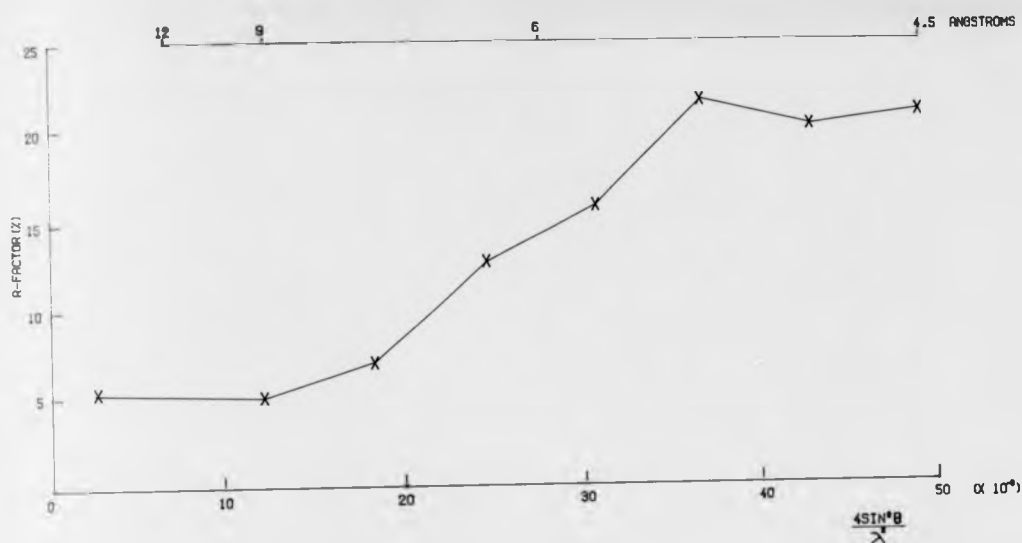


FIGURE 4.10 GRAPH SHOWING THE VARIATION OF R-FACTOR AS A FUNCTION OF RESOLUTION

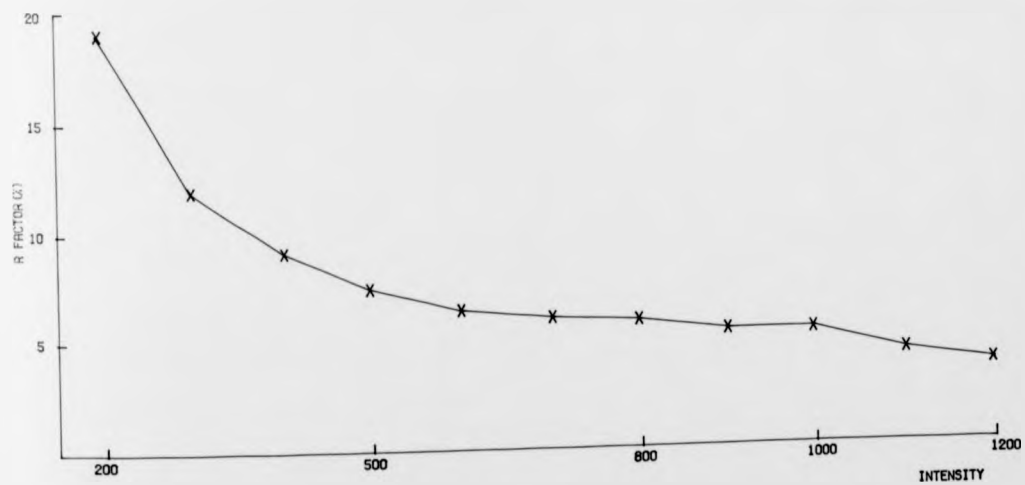


FIGURE 4.11 GRAPH SHOWING THE VARIATION OF R-FACTOR AS A FUNCTION OF INTENSITY

macromolecules in crystals is not well understood. It is dependent upon many parameters including: dose, exposure time, incident wavelength (energy), temperature, radiation protectants, sample and protein variability. It is characterised by a reduction in reflection intensity (especially at high angle) leading eventually to a total loss of high resolution data. There appear to be at least two processes involved:

- (i) The formation of free-radicals,
- (ii) Heating effects.

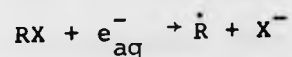
For a fuller discussion see Blundell and Johnson (1976) p.251.

(i) Free radicals

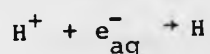
Free radicals have been detected by electron spin resonance techniques in irradiated biologically important molecules. Müller (1962) has found that for amino and nucleic acids in the dry state, in vacuum, one radical is formed per 100 eV of absorbed energy.

As protein molecules commonly contain about 50% solvent the effect of ionising radiations on water is probably the most important primary step in the radiation damage process. The chemically active radiation products of water are principally the solvated electron e_{aq}^- , the hydrogen atom H, and the hydroxyl radical $\cdot OH$. The solvated electron is a very strong reducing agent and is especially reactive

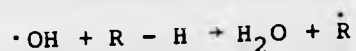
towards asymmetrical bonds causing heterolysis and the formation of free radicals:



It may also convert a hydrogen ion into a hydrogen atom:



The hydroxyl radical is a strong oxidising agent which will attack both single and double bonds, e.g:



The chain reaction initiated by free radical formation probably accounts for the common observation that radiation damage in protein crystals occur even after the X-ray beam is switched off.

The activated reaction products and free radicals caused by X-ray photons can diffuse readily through aqueous channels and inflict damage to molecules far from the site of primary absorption. It is difficult to quantify the effects of free radical damage but Blake and Phillips (1962) deduced that for each photon of CuK_{α} radiation absorbed by myoglobin 70 molecules were disrupted and a further 90 molecules somewhat disordered.

(ii) Sample lifetime

When a photon of X-radiation is absorbed it produces an excited electron. The electron may dissipate its energy in the form of heat, radiation or by direct transfer to neighbouring atoms. Any of these can lead to the destruction of a molecule in the region of the absorbed photon. Many authors have tried to quantify this effect (e.g. Helliwell and Fourme (1983), Stuhmann (1978) and Bordas (1982)). In all the models there needs to be a term to account for the susceptibility of the sample, which may be expressed in terms of the number of molecules damaged per absorbed photon. This factor must be dependent upon the parameters listed at the start of this section. Helliwell and Fourme (1983) call this factor the 'damage factor ϵ '. The lifetime of the crystal can thus be expressed:

$$\tau = \frac{N_s}{I_o d^2 \mu d \epsilon}$$

where N_s = number of molecules in a sample,

I_o = incident intensity in
photons/second/mm²/horizontal mrad,

d = size of an isodimensional crystal,

μ = absorption coefficient,

ϵ = damage factor,

$I_o d^2 \mu d$ = approximate number of photons absorbed per second.

(the exact formula is $I \cdot d^2 (1 - e^{-\mu d})$)

Various values of μ have been reported or assumed in the literature. Table 4.4 lists these values and quotes their source.

As was mentioned earlier, (see section 3.2), several workers have reported that relative crystal lifetimes (i.e. amount of data per crystal) of certain proteins have been reduced since the installation of a vertically focussing mirror at the SRS. In order to clarify the effect of the increased flux on our bacterial 6PGDH crystals, the resolution limit of each of the A-films was determined visually. This was then plotted against total exposure in the X-ray beam $\int I \cdot t$, where I is the ionisation chamber current and t is the exposure time in minutes. The results are shown in figures 4.12 and 4.13. The temperature factors for the same crystals were also plotted against $\int I \cdot t$ and are shown in figures 4.14 and 4.15.

All four crystals collected before the mirror was installed show a two phase curve. There is a plateau region where the maximum resolution remains constant (the quantum steps shown in the graphs for crystals DNB1, DNB2 and DNB4 are due to an increase in the exposure time, and can therefore be regarded as a part of the plateau). The subsequent fall off in resolution is of the order of $0.014 \text{ \AA}^{\circ} (\int I \cdot t)^{-1}$. The post mirror curves (shown in figure 4.13) do

Table 4.4 Various values of ϵ reported or assumed in the literature. Table based on Helliwell and Fourme.

ϵ	E (keV)	Source
2.5	8	Helliwell and Fourme (1983)
10	10	Stuhrmann (1978)
75	8	Blake and Phillips (1962)
168*	8	Bordas (1982)

* derived by Helliwell and Fourme (1983) on the basis of empirical formula by Bordas (1982) for calculating the maximum beam intensity for small angle work due to radiation damage.

FIGURE 4.12 GRAPHS SHOWING THE VARIATION OF RESOLUTION AS A FUNCTION OF X-RAY DOSE FOR THE PRE-MIRROR DATA

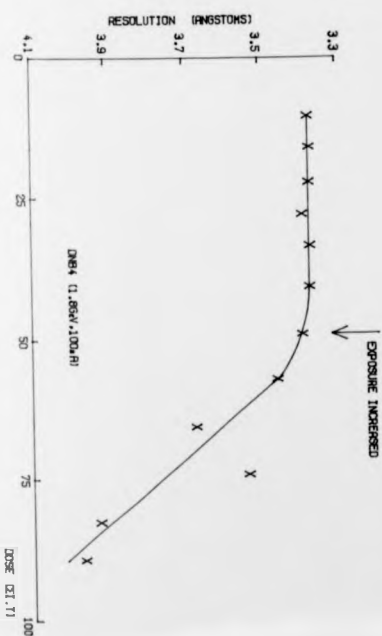
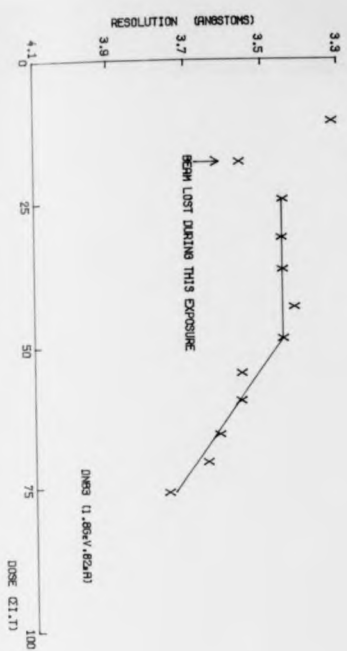
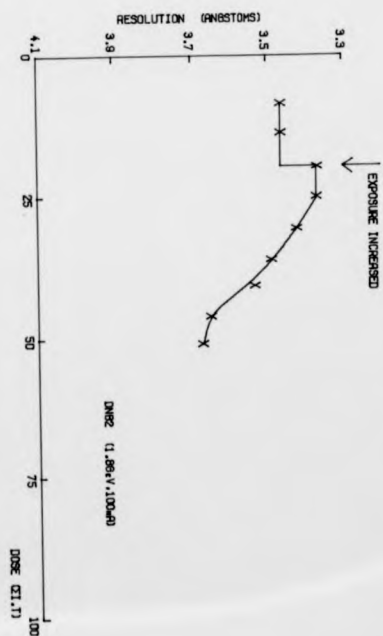
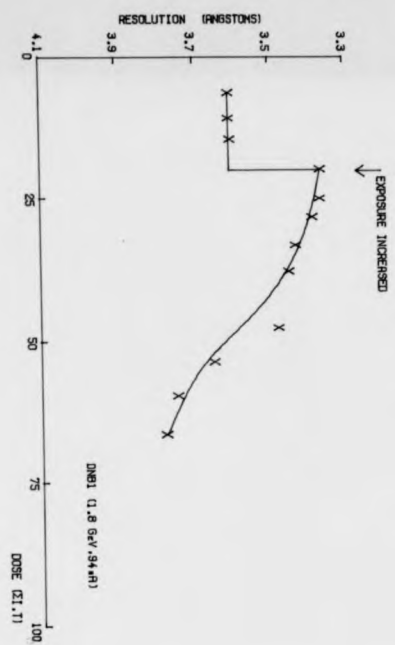


FIGURE 4.13 GRAPHS SHOWING THE VARIATION OF RESOLUTION AS A FUNCTION OF X-RAY DOSE FOR THE POST-MIRROR DATA

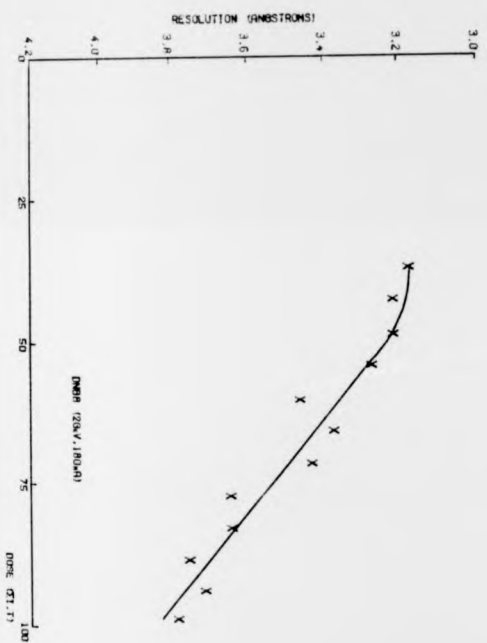
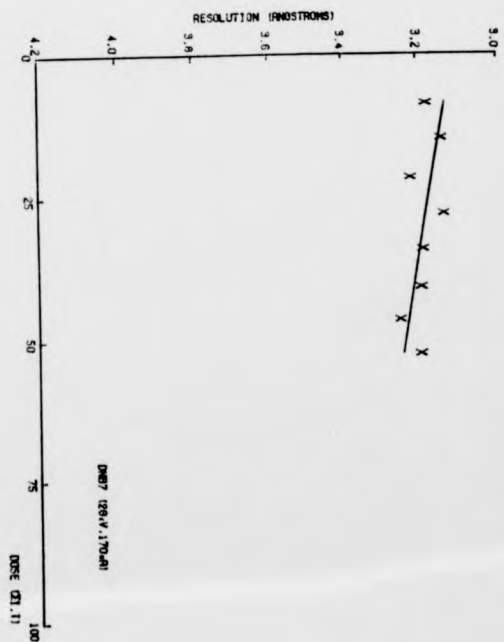
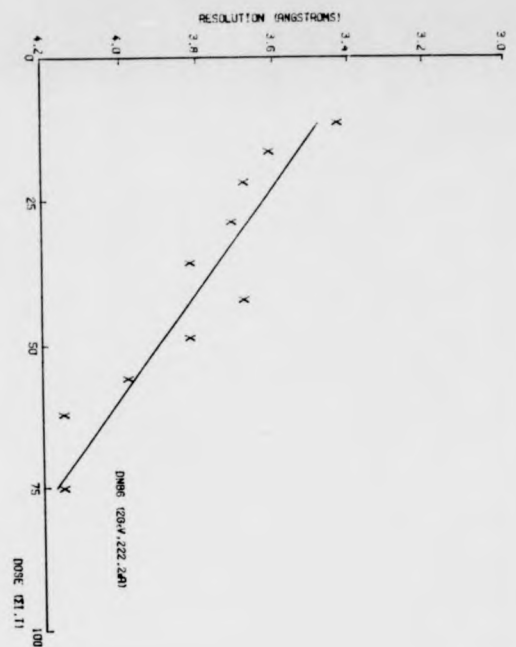
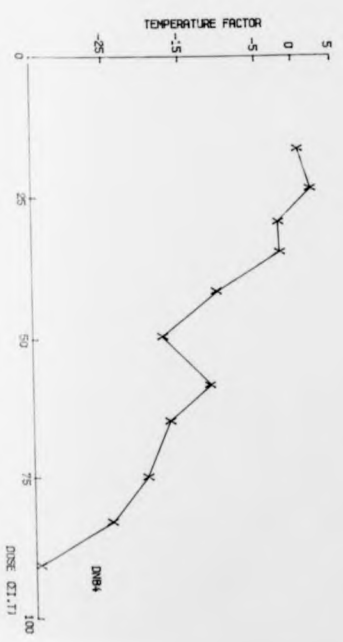
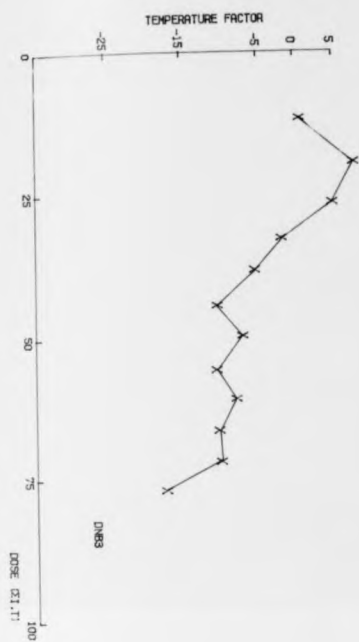
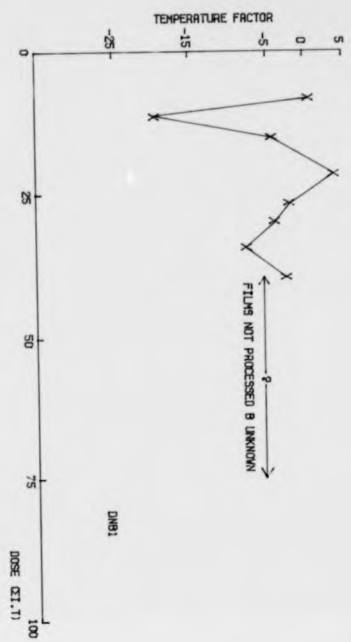


FIGURE 4.14 GRAPHS SHOWING THE VARIATION OF TEMPERATURE FACTOR AS A FUNCTION OF X-RAY DOSE FOR THE PRE-MIRROR DATA



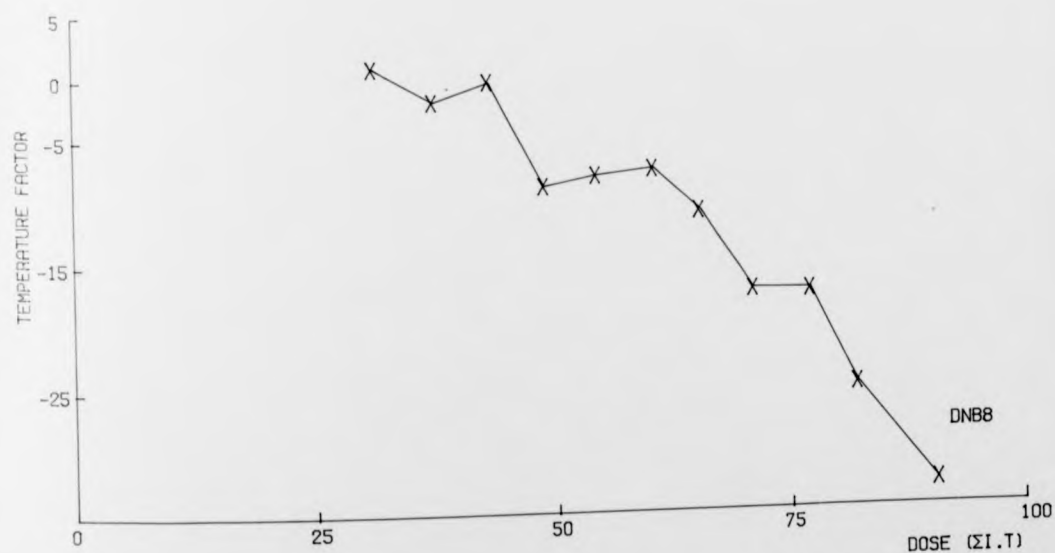
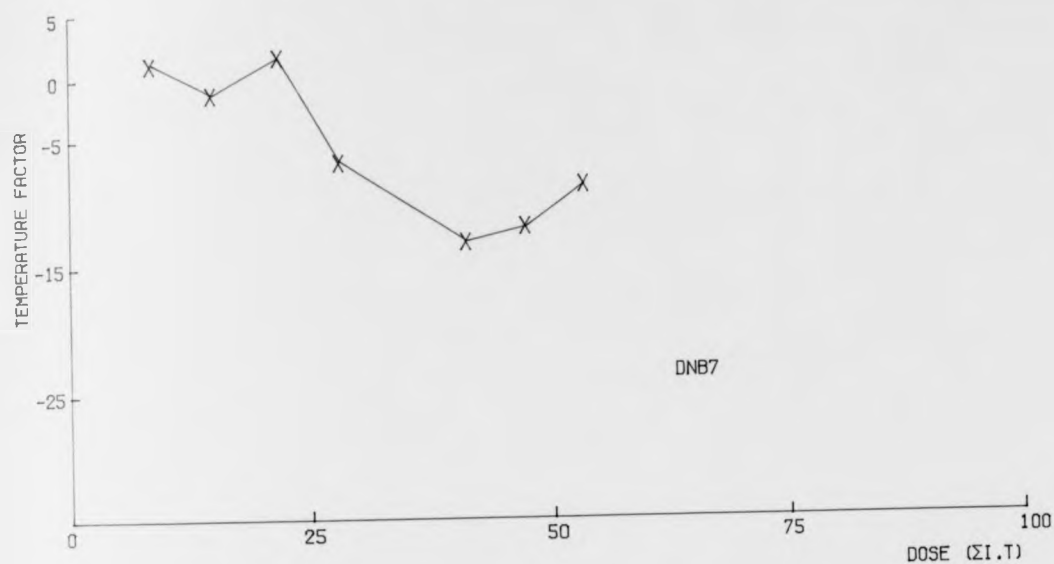


FIGURE 4.15 GRAPHS SHOWING THE VARIATION OF TEMPERATURE FACTORS AS A FUNCTION OF X-RAY DOSE FOR THE POST MIRROR DATA

not have such an obvious plateau region and tend to suffer radiation damage from the outset although the maximum resolution for crystals DNB7 and DNB8 is higher than for the pre-mirror data. The rate of decay (fall off in resolution) for crystals DNB6 and DNB8 averages to be $0.0117 \text{ \AA}^0 (\text{It}^{-1})$, i.e. very similar to the pre-mirror values. DNB7, however, decays much more slowly (approximately a factor of 12 times more slowly) but unfortunately data collection had to be terminated at the end of a shift on this crystal. It would have been interesting to see whether this crystal would continue to decay slowly or if it too had a two phase decay.

The limited data for bacterial 6PGDH radiation decay means that it would be rather foolish to try and generalise about pre and post mirror decay patterns, especially as the DNB7 curve differs so markedly from those of DNB6 and DNB8. One can, however, observe that bacterial 6PGDH does not show the very dramatic reductions in relative crystal lifetime that the groups at York and Bristol report (Dodson et al (1983), Achari et al (1983)), and indeed a higher maximum resolution was obtained with the increased flux. The B-factor plots, although showing a negative trend with dose, were not very revealing when the pre and post mirror plots were compared (see figures 4.14 and 4.15).

The lifetime of many protein crystals are increased by cooling. This presumably reduces heating effects or slows down the rate of diffusion of free radicals through the

aqueous solvent. Due to a lack of synchrotron beam time no experiments with cooling bacterial 6PGDH have been tried.

CHAPTER 5

THE ROTATION FUNCTION

5.1 Introduction

Because the heavy atom derivative (platinum) data collected at EMBL, Hamburg proved to be unprocessable (see section 3.1), it was decided to obtain initial phases from the bacterial 6PGDH by means of molecular replacement. The first set of atomic coordinates used for the molecular replacement calculations were built to the original three derivative map containing terms to 2.6 Ångstroms. (R.Pickersgill, pers. comm.). A second set of coordinates were given to us at a later date which had been built to a modified electron density map. This second set was refined against structure factor data between 10⁰Å and 2.6⁰Å. The R-factor was 0.418. Section 5.4.1 gives the results using this dataset.

The history and theory of molecular replacement are well described in a collection of papers edited by M.G. Rossmann (1972). A brief outline of the theory is given below. The Rotation function was used to try to locate the non-crystallographic two-fold using the intensity measurements from the bacterial data. It was also used to

try to fit the sheep liver atomic coordinates into the bacterial cell.

5.2 Rotation Function Theory

The Fourier transform of the square of the structure factor amplitudes is called the Patterson function and is used extensively in molecular replacement. The fact that it is the transform of $|F|^2$ means that it can be calculated directly from observed intensities with no knowledge of phases. The Patterson function contains peaks at positions relative to the origin of all the interatomic vectors with peak heights proportional to the product of the atomic numbers.

For very small molecules structure determination can be accomplished by means of the Patterson, but it is far too complicated, and the peaks are far too badly resolved for direct use in protein crystallography. However, Rossmann and Blow (1962) show that by using an overlap function and carefully defining a sphere of integration in Patterson space along with an interference function $G_{\underline{h}\underline{h}}$, the relative orientations of identical sub-units can be detected. The Rotation (overlap) function defined by them is:

$$R = \frac{U}{V^3} \sum_{\underline{h}} \sum_{\underline{p}} |F_{\underline{h}}|^2 |F_{\underline{p}}|^2 G_{\underline{h}\underline{h}},$$

where U = volume of Patterson space used,
 V = volume of unit cell,
 \underline{h} = reciprocal lattice vector,
 \underline{p} = reciprocal lattice vector,
 $G_{\underline{h}\underline{h}'}$ = interference function,
 \underline{h} = non integral reciprocal lattice
 vector given by $\underline{p} = [\hat{C}] \underline{h}$,
 and $[\hat{C}]$ is the transpose of $[C]$ which is the
 rotation matrix required to
 superimpose the Patterson peaks.

Tollin and Rossmann (1966) showed that because of the
 nature of the interference function, the Rotation function
 could more conveniently be written:

$$R = \frac{U}{V^3} \sum_{\underline{p}} |F_{\underline{p}}|^2 \left\{ \sum_{\underline{h}} |F_{\underline{h}}|^2 G_{\underline{h}\underline{h}'} \right\}$$

The inner summation $\left\{ \sum_{\underline{h}} |F_{\underline{h}}|^2 G_{\underline{h}\underline{h}'} \right\}$ need only be made for a
 small number of points close to the non-integral vector $-\underline{h}'$.
 Tollin and Rossmann found that using as few as 27 points was
 satisfactory.

The calculations are accelerated by using Eulerian
 coordinates and Crowther's (1971) FFT method of calculating
 the Patterson density in terms of spherical harmonics rather
 than Cartesian Fourier components $|F|^2$.

The Rotation function can be used both to detect non crystallographic symmetry between sub-units (i.e. $|F_h|^2$ is a rotated Patterson which is superimposed on $|F_p|^2$ the non-rotated Patterson), or to fit a 'known structure' into the cell of an 'unknown structure' (i.e. $|F_h|^2$ is the Patterson of the known structure and is rotated onto the Patterson $|F_p|^2$ of the unsolved structure obtained from intensity measurements). Figure 5.1 shows how careful selection of the sphere of integration separates the 'self' and 'cross' Patterson vectors. There is a mirror plane shown along DC in the Patterson. The real space atoms do not have any mirror symmetry which illustrates that the symmetry of the Patterson is not the same as the real space symmetry. A full discussion of rotation space symmetry is given by Tollin, Main and Rossmann (1966) and the space groups for all non cubic rotation functions are calculated by Rao et al (1980).

5.3 Bacterial 6PGDH Self-Rotations

The observed structure factor amplitudes from the bacterial 6PGDH data were input into the program ALMN. ALMN is a version of Crowther's FFT rotation function (Crowther (1971)). It has been extensively improved by Eleanor Dodson who has extended the output to detail symmetry related Eulerian angles, spherical polars and directional cosines for the maximum value of R (rotation function) on each β

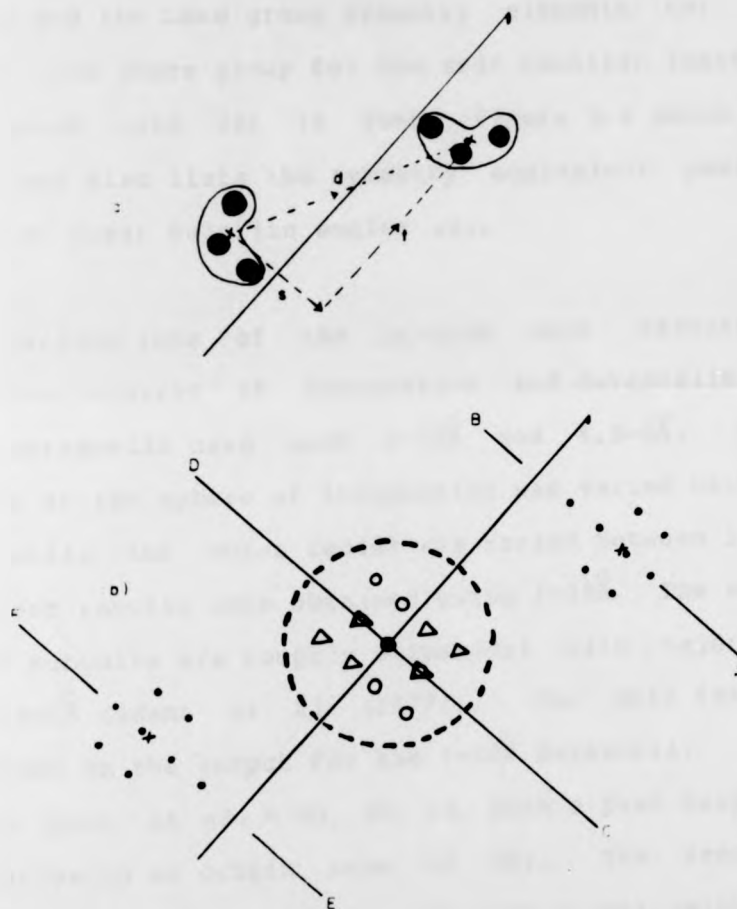


Figure 5.1 The Patterson of a structure with a screw rotation axis. (a) represents atoms in a two-dimensional model structure. The crosses are the points chosen as the approximate centres of the two molecules which are related by a vector d . d has components t and s parallel and perpendicular to the screw rotation axis. (b) shows the vectors arising from the structure in (a). The self vectors of the two molecules are indicated by Δ and \circ ; the cross vectors by \bullet . Crosses mark the positions of d and $-d$. CD is a mirror plane through the origin; AB and EF are Harker planes.

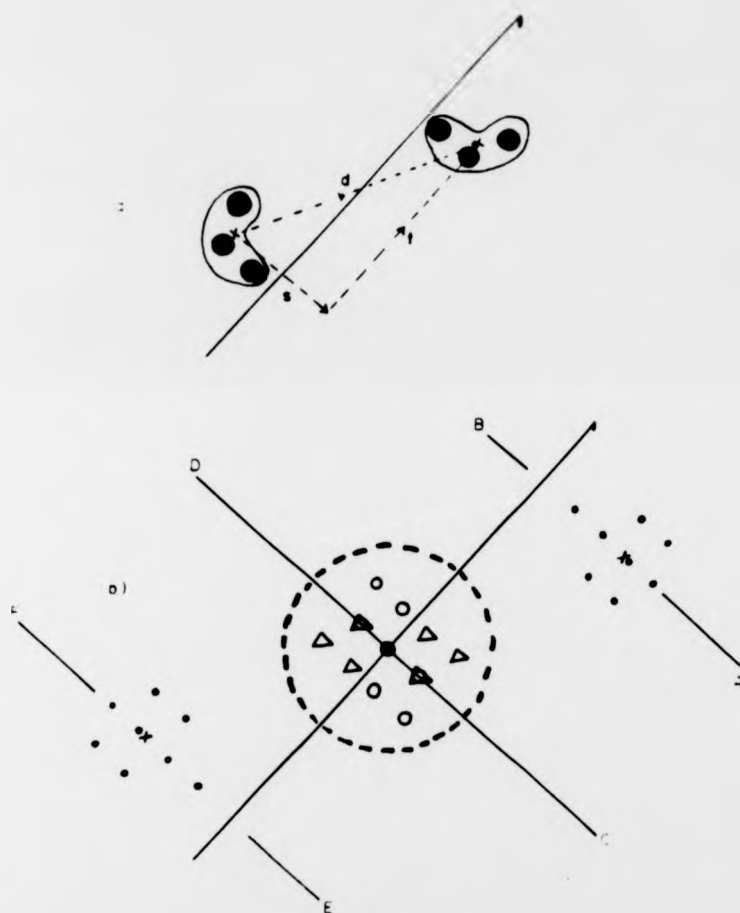


Figure 5.1 The Patterson of a structure with a screw rotation axis. (a) represents atoms in a two-dimensional model structure. The crosses are the points chosen as the approximate centres of the two molecules which are related by a vector d . d has components t and s parallel and perpendicular to the screw rotation axis. (b) shows the vectors arising from the structure in (a). The self vectors of the two molecules are indicated by Δ and ∇ ; the cross vectors by \odot . Crosses mark the positions of d and $-d$. CD is a mirror plane through the origin. AB and EF are Harker planes.

section of the output Eulerian angle map. The relationships between Eulerian angles $\alpha\beta\gamma$, the spherical polars $\omega\phi\chi$ and the directional cosines DC1, DC2, DC3 are shown in figure 5.2. Figure 5.3 shows the relationship between the axial system and the Laue group symmetry elements for bacterial 6PGDH. The space group for the self rotation function (i.e. $\bar{3}m$ rotated onto $\bar{3}m$) is Pbnb. Figure 5.4 shows the space group and also lists the symmetry equivalent positions in terms of their Eulerian angles $\alpha\beta\gamma$.

Various runs of the program were executed using different spheres of integration and datashells. The two main datashells used were 6-10 \AA and 4.5-6 \AA . The inner radius of the sphere of integration was varied between 3 and 10 \AA , while the outer radius was varied between 10 and 40 \AA . The best results were obtained using 3-30 \AA . The sheep liver 6PGDH subunits are roughly ellipsoidal with major axes of 90x60x60 \AA (Adams et al (1977)). Two main features were observed on the output for the 6-10 \AA datashell; the first was a peak at $\alpha\beta\gamma = 90, 30, 90$, with a peak height of 16.8 (relative to an origin peak of 50). The second was a smaller peak at $\alpha\beta\gamma = 30, 20, 30$ with a peak height of 13.2. Both of these peaks were fairly broad with density spreading over several 5 $^\circ$ steps in Eulerian space.

When the 4.5-6 \AA datashell was used, the second of these two features was not observed, the original peak at $\alpha\beta\gamma = 90, 30, 90$ being the only major feature of the output

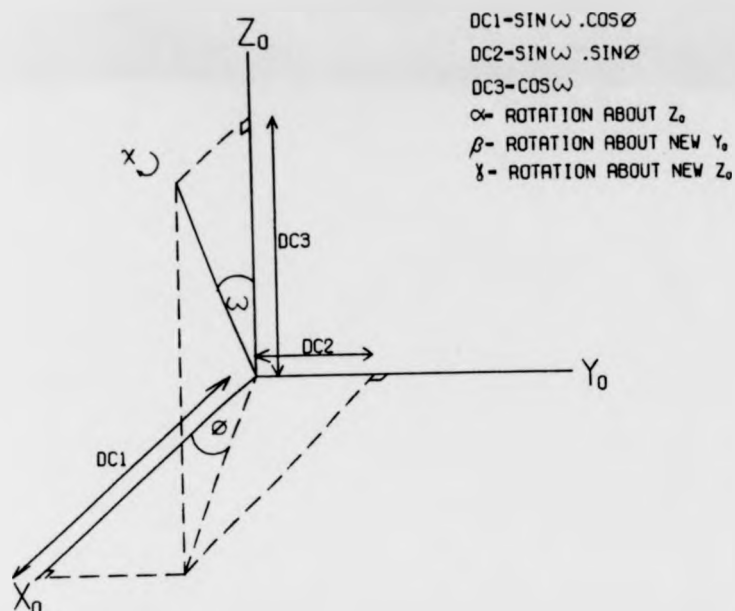


FIGURE 5.2 THE RELATIONSHIP BETWEEN THE CARTESIAN COORDINATE FRAME, THE SPHERICAL POLARS AND THE DIRECTIONAL COSINES

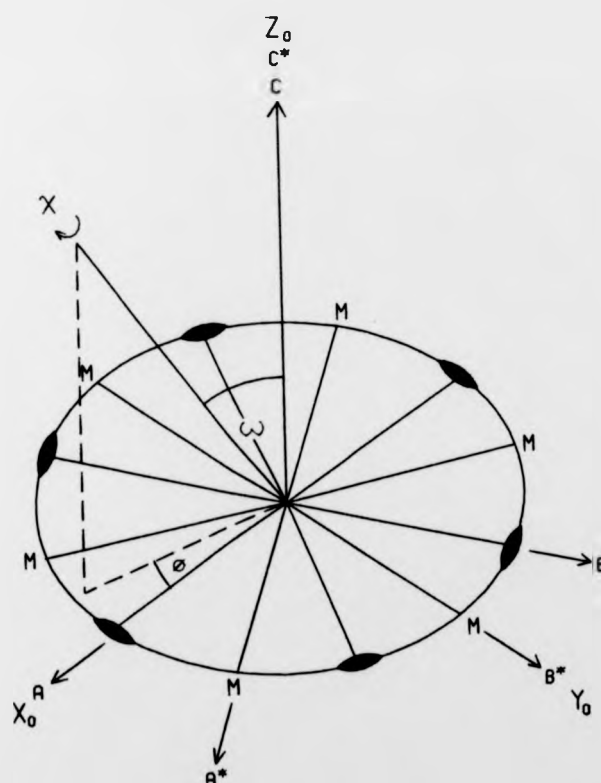


FIGURE 5.3 THE RELATIONSHIP BETWEEN THE AXIAL SYSTEM, LAUE GROUP AND UNIT CELL DIRECTIONS

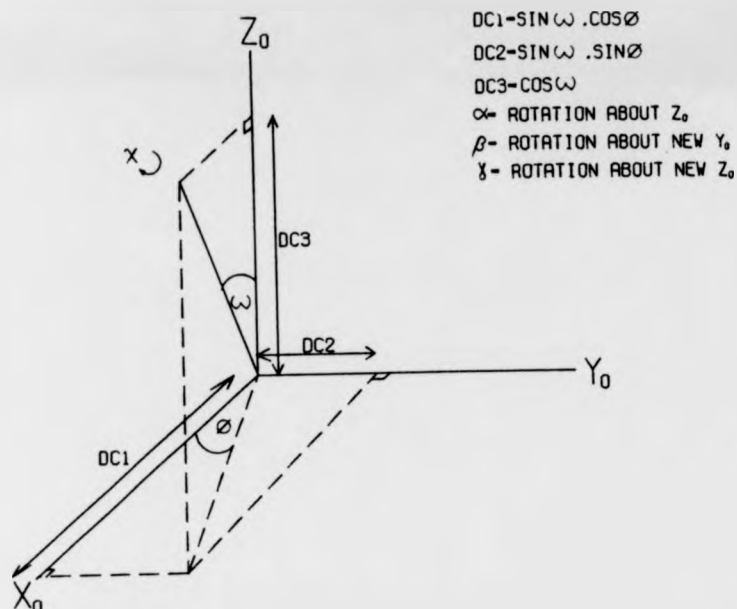


FIGURE 5.2 THE RELATIONSHIP BETWEEN THE CARTESIAN COORDINATE FRAME, THE SPHERICAL POLARS AND THE DIRECTIONAL COSINES

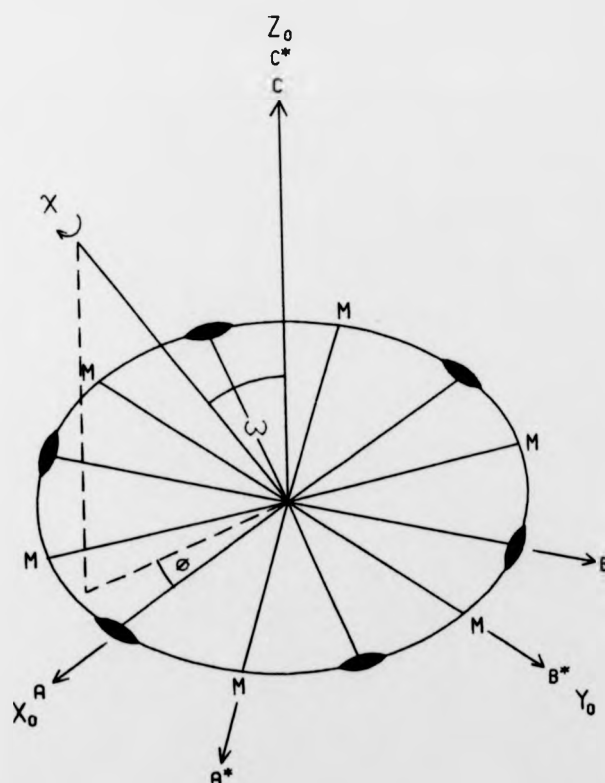


FIGURE 5.3 THE RELATIONSHIP BETWEEN THE AXIAL SYSTEM, LAUE GROUP AND UNIT CELL DIRECTIONS

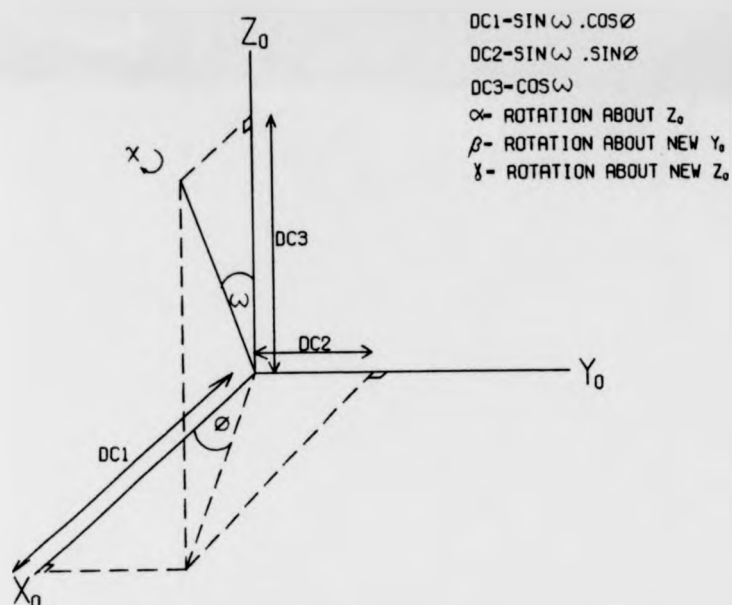


FIGURE 5.2 THE RELATIONSHIP BETWEEN THE CARTESIAN COORDINATE FRAME, THE SPHERICAL POLARS AND THE DIRECTIONAL COSINES

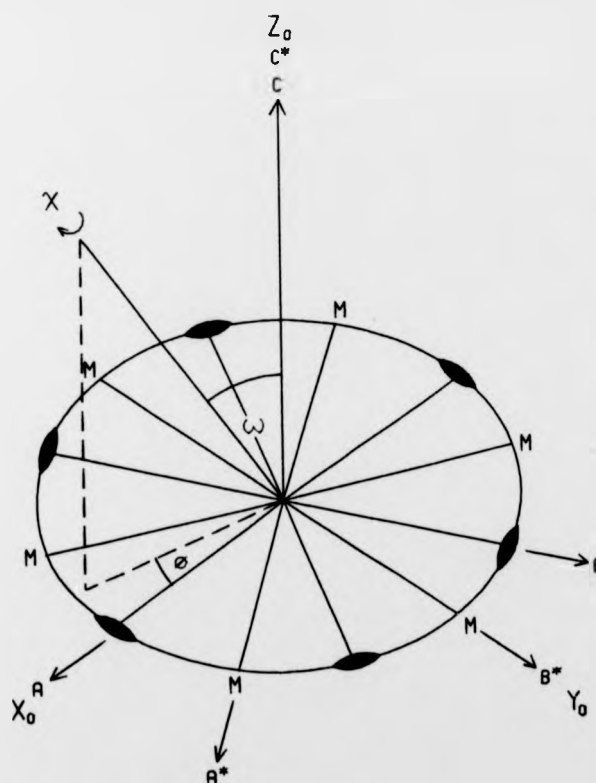


FIGURE 5.3 THE RELATIONSHIP BETWEEN THE AXIAL SYSTEM, LAUE GROUP AND UNIT CELL DIRECTIONS



- FIGURE 5.4 SYMMETRY OF THE SELF ROTATION FUNCTION
($\bar{3}_m$ ROTATED ON $\bar{3}_m$ PRODUCES P_{60})

Eulerian map. The peak was sharper for this datashell than for the 6-10 \AA data and had a peak height of 11. The peak heights produced by runs using different data shells are not directly comparable because the program scales all the output values of R so that the origin peak has a value of 50. Changing the size of the integration sphere changed the relative peak heights but did not alter the general features of the self rotation maps. Contoured maps showing both the peaks are given in figures 5.5 and 5.6.

For both of these peaks the spherical polar ϕ value is perpendicular to one of the crystallographic two-fold axes. This leads to an ambiguity in the ω value. Figure 5.7 shows the cause of this ambiguity. The crystallographic two-fold generates a second dimer $A'B'$ from the original dimer AB , therefore there are two non-crystallographic two-fold axes present, the first of these is the molecular dimer axis relating molecule A to molecule B and A' to B' . The second non-crystallographic two-fold axis is at 90° to the first and relates A to B' and B to A' . The program ALMN cannot distinguish between these two and lists both values as being possible. Consequently we have two possibilities for the spherical polars for these peaks. They are $\omega \phi \chi = 10, 30, 180$ or $100, 30, 180$ for the $\alpha\beta\gamma = 30, 20, 30$ peak and $\omega \phi \chi = 15, 90, 180$ or $105, 90, 180$ for the original $\alpha\beta\gamma = 90, 30, 90$ peak. Clearly we wish to pick the one corresponding to the close packing dimer.

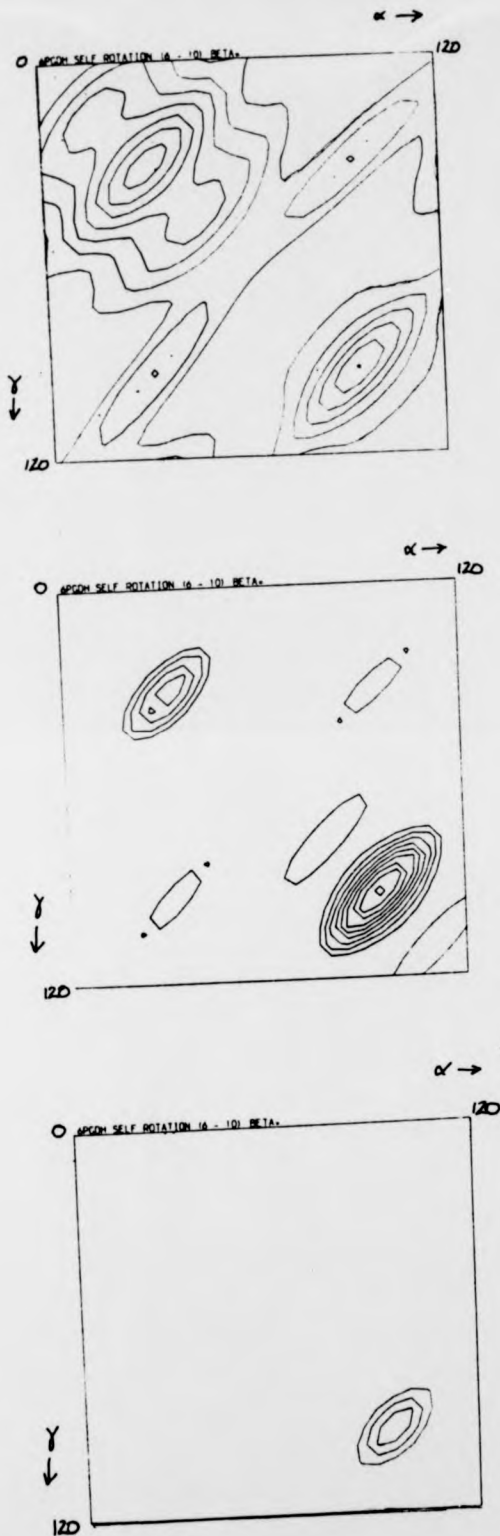


Figure 5.5 Contour Maps of the Self Rotation Function
Using the 6-10 Å Datasheet. Contours every 1.0 above 7.0

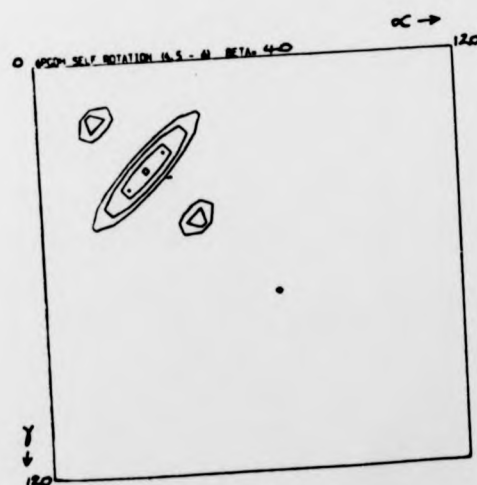
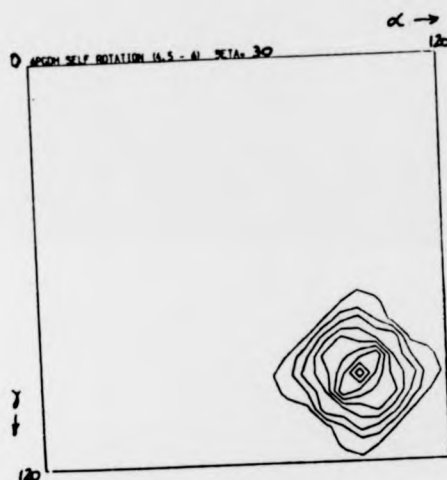
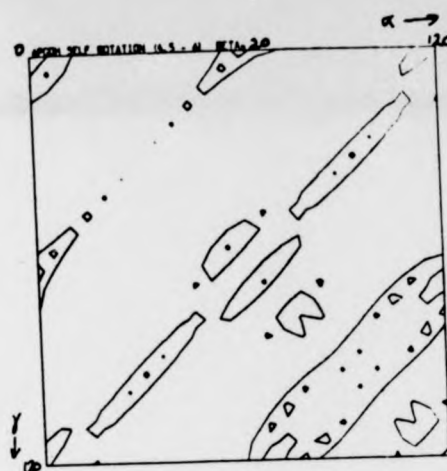


Figure 5.6 Contour Maps of the Self Rotation Function
Using the 4.5-6 Å Datashell. Contours every 1.0 above 2.0

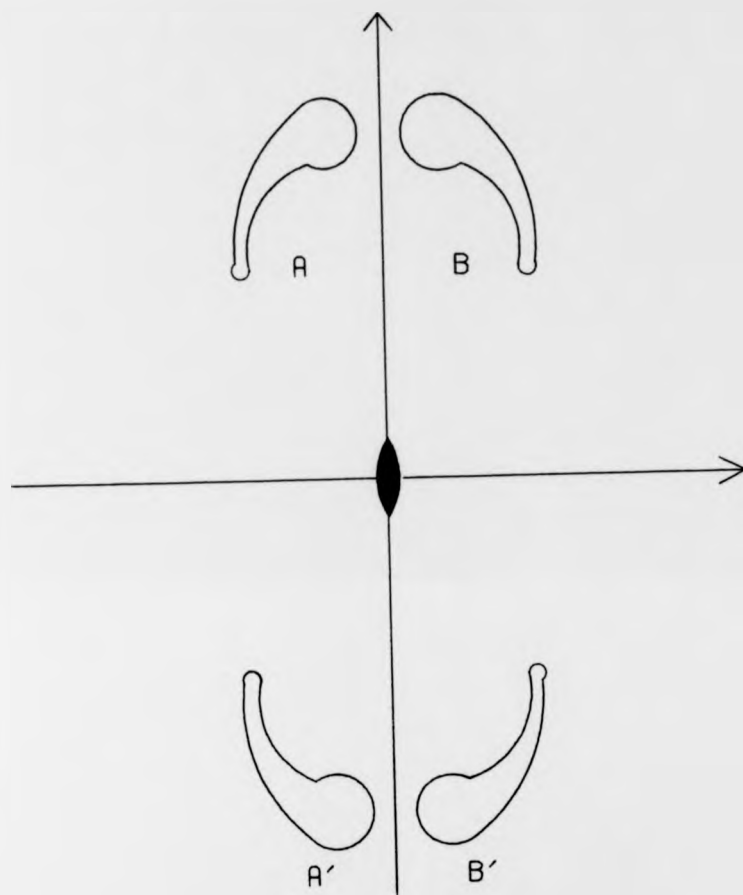


FIGURE 5.7 DIAGRAM TO SHOW THE TWO NON CRYSTALLOGRAPHIC TWO-FOLD AXES PRODUCED WHEN THE DIMER AXIS IS PERPENDICULAR TO A CRYSTALLOGRAPHIC TWO FOLD (• REPRESENTS THE SECOND NON-CRYSTALLOGRAPHIC TWO-FOLD OUT OF THE PLANE OF THE PAPER)

The fact that in the sheep liver cell the dimer axis is parallel to the longest cell axis ($b = 148.15\text{\AA}$) and that for the $\omega\phi\chi = 15, 90, 180$ case the dimer axis is at 15° to the longest axis ($c = 147.7\text{\AA}$) in the bacterial cell would seem to indicate that this orientation is more probable than $\omega\phi\chi = 105, 90, 180$, especially as the two unit cell axes have the same length to within 0.5\AA .

In summary the self rotation runs show there to be a non-crystallographic dimer axis at either $\omega\phi\chi = 15, 90, 180$ or $105, 90, 180$ in the bacterial 6PGDH cell. This peak on the rotation maps persisted when one changed the data used to obtain the initial $|F|^2$ and when one changed the size of the Patterson sphere of integration. The size of the unit cell would possibly seem to favour the first of these two.

5.4 Bacterial/Sheep Liver 6PGDH Cross Rotation Functions

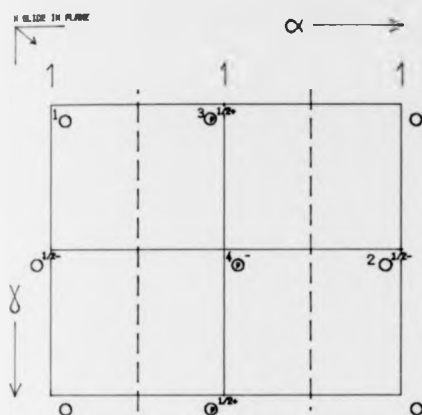
The sheep liver 6PGDH crystallizes in the space group $C222_1$ with unit cell parameters 72.72, 148.15, 102.91, 90, 90, 90. There is one monomer in the asymmetric unit and the dimer axis is crystallographic at $z = 1/4$, the two monomers being related by the symmetry operator $-x, y, 1/2 - z$. The bacterial 6PGDH crystallizes in the space group $P3_121$ (or $P3_221$) with unit cell parameters 123.1, 123.1, 147.7, 90, 90, 120. There is a dimer in the asymmetric unit and hence the dimer axis is non crystallographic

The sheep liver atomic coordinates were kindly sent to us by the Oxford group. These coordinates were used to calculate structure factors using the program GENSF. Two sets of structure factors were calculated; the first was for a monomer in a P1 cell. The size of the cell used was actually that of the sheep liver cell but with only one monomer present, and xyz symmetry only. Rotation function runs using the Patterson from these structure factors rotated onto the observed bacterial $|F|^2$ will henceforth be called 'monomer cross' rotation functions.

The structure factors for the second monomer were created using the symmetry operator $-x, y, 1/2 - z$ and the calculated structure factors for the first monomer using the program CAD. The output from CAD was in the form $F_{1\alpha_1}, F_{2\alpha_2}$ for each hkl, where $F_{1\alpha_1}$ are the amplitude and phase of the structure factors from the first monomer, and $F_{2\alpha_2}$ are the amplitude and phase of the structure factors from the second monomer. A program was written to obtain the vector sum of these two. The rotation function runs using the Patterson from these structure factors rotated onto the observed $|F|^2$ will henceforth be called 'dimer cross' rotation functions.

The rotation function space groups for the monomer cross and dimer cross rotation functions are shown in figures 5.8 and 5.9.

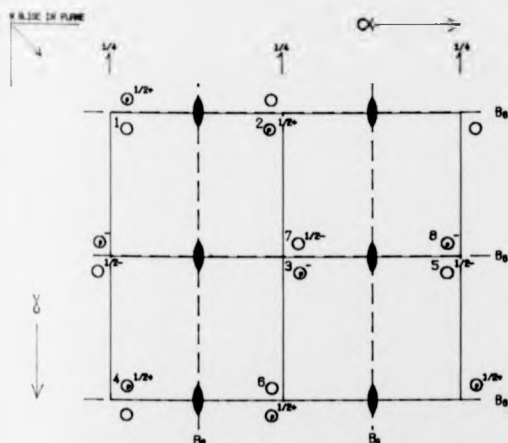
We know that the dimer axis in the sheep liver cell is



UNIT CELL - 120 IN α , 360 IN β , 360 IN γ

1. α, β, γ
2. $120-\alpha, 180-\beta, 180+\gamma$
3. $60-\alpha, 180+\beta, \gamma$
4. $60+\alpha, 360-\beta, 180+\gamma$

FIGURE 5.8 SYMMETRY OF THE MONOMER CROSS ROTATION FUNCTION (P1 ROTATED ON $\bar{3}M$ PRODUCES P_{R12_1})



UNIT CELL - 120 IN α , 360 IN β , 360 IN γ

1. α, β, γ
2. $60-\alpha, 180+\beta, \gamma$
3. $60+\alpha, 360-\beta, 180+\gamma$
4. $\alpha, 180+\beta, 360-\gamma$
5. $120-\alpha, 180-\beta, 180+\gamma$
6. $60-\alpha, \beta, 360-\gamma$
7. $60+\alpha, 180-\beta, 180-\gamma$
8. $120-\alpha, 360-\beta, 180-\gamma$

FIGURE 5.9 SYMMETRY OF THE DIMER CROSS ROTATION FUNCTION (2/M ROTATED ON $\bar{3}M$ PRODUCES P_{R12_2} . NOTE THE ORIGIN SHIFT COMPARED TO FIGURE 5.4)

parallel to the b axis and the self rotation tells us that the dimer axis lies at an angle of 15° to the c axis in the bacterial cell (or 105° to c). If we define our axes so that for the bacterial cell c is along Z_0 (see figure 5.2) and for the sheep liver cell b is along Z_0 , then we would expect that the cross rotation will peak at an $\omega\phi$ value of 15, 90 (or 105, 90). There should not be peaks at both 15, 90 and 105, 90. We can hopefully use this knowledge to resolve the ambiguity of the self-rotation result. The correct solution will also give us a value for χ so that the orientation of the molecule with respect to the $P3_21$ cell can be fully described. The self rotation was unable to yield the χ value because for a two fold axis χ is necessarily equal to 180° for a self rotation.

Unfortunately cross rotations are notorious for giving false peaks and showing alarming variation in peak position and height when the calculating parameters are changed slightly (e.g. Ru-chang et al (1983), Smith et al (1983)).

It was found that better results were obtained when the original $|F|$'s were sharpened by multiplying them by $\exp(-B \sin^2\theta / \lambda^2)$. A sharpened Patterson has vectors corresponding to the distances between point atoms rather than atoms with a finite sphere of electron density. B values of -20 gave the best results in our experience.

From the monomer cross rotation function the $6-10\text{\AA}$

datashell produced 22 peaks with peak heights ranging from 73 to 143. The 4.5-6 \AA datashell produced 35 peaks with peak heights ranging from 67-111. A Patterson sphere of 3-30 \AA was used for both. Table 5.1 shows all 57 peaks with the rotation function value (relative to an origin of 50) produced from both datashells. The peaks marked by an asterisk were considered reasonable but the unmarked peaks were rejected as inconsistent.

The dimer cross rotation function for the 6-10 \AA datashell produced 20 peaks with peak heights ranging from 85 to 154. The 4.5-6 \AA datashell produced 16 peaks with peak heights ranging from 81 to 99. A composite table showing the 'dimer cross' peaks along with the value of the rotation function for both datashells is given in table 5.2. Again the most hopeful peaks are marked by asterisks.

There were only three peaks out of all the cross rotation peaks that showed any kind of consistency for runs of both the dimer and monomer, and for both datashells. These are marked by + in table 5.2 and have Eulerian angles $\alpha\beta\gamma = (35,45,345)$, $(55,60,30)$ and $(110,80,125)$. The peaks were not at exactly these positions on all the maps but there were peaks or regions of high density very close to these coordinates.

The program ALMN was used to calculate the spherical polars for each of these three peaks and all their symmetry

Table 5.1 Cross-rotation peaks obtained using structure factors from the sheep liver monomer atomic coordinates in a P1 cell and the observed bacterial structure factors.

a	B	γ	RF ₆₋₁₀	RF _{4.5-6}	
5	10	30	84	85	*
45	10	85	-12	81	
90	10	55	55	77	*
110	10	40	89	75	*
30	15	40	22	96	*
95	15	145	16	73	*
75	15	155	0	67	
115	15	35	120	10	
0	15	160	121	1	
110	15	170	128	0	
95	15	300	113	-39	
50	25	55	-22	75	
95	25	65	16	77	*
75	25	20	107	19	*
55	25	335	97	10	
20	30	175	73	11	
110	30	45	44	82	*
45	30	175	58	71	*
80	35	25	74	80	*
65	35	140	-25	77	
110	35	15	113	-13	
85	40	45	109	35	*
100	40	55	1141	0	
70	40	265	79	-40	
90	40	125	15	80	
115	40	25	69	111	*
80	45	90	-5	98	
60	45	180	-33	84	
85	45	190	-46	90	
30	45	250	-18	80	
85	45	40	107	60	*
40	50	165	83	34	*
65	50	5	111	6	
5	50	25	75	95	*
80	50	40	75	91	*
90	50	45	49	92	*
85	50	245	44	96	*
70	55	55	-12	105	
10	55	35	47	99	*
95	55	355	97	-30	
15	60	20	63	70	*
40	60	150	0	105	
55	60	15	108	-5	*
45	65	60	37	96	*
75	65	310	-100	101	
115	70	65	50	105	*
85	70	15	118	37	*
15	70	105	135	-52	
20	75	285	-70	105	
55	75	280	-64	100	
50	80	300	-62	101	
105	80	115	143	42	*
105	80	350	92	-32	
65	85	15	115	10	
110	85	115	88	107	*
50	90	195	104	-31	
85	90	70	-71	110	

+ 56 @ 50 60 10
+ 98 @ 50 60 10

Table 5.2 Cross-rotation peaks obtained using structure factors from the sheep liver dimer atomic coordinates in a P1 cell and the observed bacterial structure factors

α	β	γ	RF ₆₋₁₀	RF _{4.5-6}	
0	15	160	99	-25	
110	20	175	112	3	
85	20	75	-15	85	
90	20	290	-6	85	
10	25	25	103	44	*
100	30	55	0	88	
105	30	305	23	80	*
75	30	310	5	88	
15	35	180	113	50	*
75	40	55	107	5	
35	45	170	-40	99	
35	45	345	39	93	*
15	50	10	18	98	*+
40	50	170	-39	94	
65	50	90	85	-56	
80	55	10	109	17	*
30	55	25	44	81	*
55	60	30	70	85	*
50	60	10	100	44	*+
70	60	115	138	-58	
0	65	240	111	11	
115	65	240	122	2	
85	70	80	-63	80	
65	75	145	128	-5	
10	80	0	135	-49	
0	80	215	123	-4	
65	80	315	134	-22	
40	80	55	52	81	*
110	80	125	61	85	*+
75	85	10	154	-31	
60	85	20	134	-1	
45	90	155	133	16	*
15	90	170	138	-32	
75	90	175	110	-44	
50	90	40	39	90	*
75	90	45	53	81	*

equivalents. Table 5.3 shows the results. The program calculates 6 symmetry equivalents (c.f. 4 symmetry equivalents shown in figure 5.8), the additional two can be derived by means of unit cell translations on the symmetry elements given in figure 5.8. The last column of table 5.3 shows which of the symmetry elements from figure 5.8 were used and the appropriate unit cell translations applied. The reason for this apparent redundancy of information is that ALMN fully expands the symmetry equivalents and includes those which only require unit cell translations to bring them into the unit cell of the rotation function space group.

However, none of the three peaks or their symmetry equivalents gave $\omega\phi$ values of 15, 90 or 105, 90 (or Laue group equivalents). Neither did they give $\omega\phi$ values of 10, 30, or 100, 30 (or Laue group equivalents), which would correspond to the smaller self rotation feature. Because a simple inspection of the major peaks did not reveal the information we were expecting from the self rotation results, we took a closer look at where we would expect to see maxima on the cross rotation maps.

We know that for the correct peak the rotation matrix $[\Omega]$ will satisfy the condition:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{P3_221} = [\Omega] \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{C222_1}$$

Table 5.3 Spherical polars and equivalents of the three 'best' peaks (see text)

α	β	γ	Equivalent	ω	ϕ	χ	Symmetry operator
35	45	-15	α, β, γ	67.3	115	49.0	(1)
155	45	-15	$120+\alpha, \beta, \gamma$	23.8	175	143.2	(1) + 1 unit cell in α
95	-45	165	$60+\alpha, -\beta, 180+\gamma$	151.6	55	107.1	(4)
85	135	165	$120-\alpha, 180-\beta, 180+\gamma$	108.7	-130	154.6	(2)
25	-135	-15	$60-\alpha, -180+\beta, \gamma$	87.9	-70	135.2	(3) - 1 unit cell in β
145	-135	-15	$180-\alpha, \beta-180, \gamma$	69.4	-10	161.4	(3) + 1 unit cell in α ; -1 unit cell in β
55	60	30	α, β, γ	40.5	102.5	100.6	(1)
175	60	30	$120+\alpha, \beta, \gamma$	149.41	-17.5	158.4	(1) + 1 unit cell in α
115	-60	-150	$60+\alpha, -\beta, \gamma-180$	117.5	42.5	68.6	(4) - 1 unit cell in γ
65	120	-150	$120-\alpha, 180-\beta, \gamma-180$	111.3	-162.5	136.7	(2) - 1 unit cell in γ
5	-120	30	$60-\alpha, \beta-180, \gamma$	80.2	-102.5	123.0	(3) - 1 unit cell in β
125	-120	30	$180-\alpha, \beta-180, \gamma$	60.6	-42.5	167.6	(3) + 1 unit cell in α ; -1 unit cell in β
110	80	125	α, β, γ	136.6	-97.5	138.6	(1)
50	-80	-55	$\alpha-60, -\beta, \gamma-180$	93.0	-37.5	80.1	(4) - 1 unit cell in α ; -1 unit cell in γ
170	-80	-55	$\alpha+60, -\beta, \gamma-180$	44.9	22.5	131.4	(4) - 1 unit cell in γ
10	100	-55	$120-\alpha, 180-\beta, \gamma-180$	107.8	122.5	107.1	(2) - unit cell in γ
130	100	-55	$240-\alpha, 180-\beta, \gamma-180$	62.9	-177.5	118.7	(2) + 1 unit cell in α ; -1 unit cell in γ
70	-100	125	$180-\alpha, -180+\beta, \gamma$	129.8	62.5	170.4	(3) + 1 unit cell in α ; -1 unit cell in β

The rotation matrix $[\Omega]$ can be expressed in terms of the Eulerian angles $\alpha\beta\gamma$;

$$[\Omega] = \begin{bmatrix} \cos\alpha.\cos\beta.\cos\gamma & -\cos\alpha.\cos\beta.\sin\gamma & \cos\alpha.\sin\beta \\ -\sin\alpha.\sin\gamma & -\sin\alpha.\cos\gamma & \\ \sin\alpha.\cos\beta.\cos\gamma & -\sin\alpha.\cos\beta.\sin\gamma & \sin\alpha.\sin\beta \\ +\cos\alpha.\sin\gamma & +\cos\alpha.\cos\gamma & \\ -\sin\beta.\cos\gamma & \sin\beta.\sin\gamma & \cos\beta \end{bmatrix}$$

or in terms of a rotation angle ψ about an axis whose direction is given by the directional cosines DC1, DC2, DC3:

$$[\Omega] = \begin{bmatrix} \cos\psi + DC1^2(1-\cos\psi) & DC1.DC2(1-\cos\psi) - DC3(\sin\psi) & DC3.DC1(1-\cos\psi) + DC2(\sin\psi) \\ DC1.DC2(1-\cos\psi) + DC3(\sin\psi) & \cos\psi + DC2^2(1-\cos\psi) & DC2.DC3(1-\cos\psi) - DC1\sin\psi \\ DC3.DC1(1-\cos\psi) - DC2(\sin\psi) & DC2.DC3(1-\cos\psi) + DC1(\sin\psi) & \cos\psi + DC3^2(1-\cos\psi) \end{bmatrix}$$

DC1, DC2 and DC3 can be expressed in terms of the spherical polars $\omega\phi\gamma$ by the following expressions:

$$DC1 = \sin\omega. \cos\phi$$

$$DC2 = \sin\omega. \sin\phi$$

$$DC3 = \cos\omega$$

(see figure 5.2).

Now we know the possible values of ω and ϕ from the self rotation result, and so we can deduce that:

$$[\Omega] \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} C222_1 = \begin{bmatrix} \sin\omega \cdot \cos\phi \\ \sin\omega \cdot \sin\phi \\ \cos\omega \end{bmatrix} P3_2 21$$

i.e. that the rotation matrix acting on an axis parallel to b in the C222₁ cell will rotate the axis so that its directional cosines will be as given above in the P3₂21 cell.

By substituting the Ω_{12} , Ω_{22} , Ω_{32} elements from the Eulerian angle expression and comparing matrix elements we obtain:

$$-\cos\alpha \cdot \cos\beta \cdot \sin\gamma - \sin\alpha \cdot \cos\gamma = \sin\omega \cdot \cos\phi \quad (1)$$

$$-\sin\alpha \cdot \cos\beta \cdot \cos\gamma + \cos\alpha \cdot \cos\gamma = \sin\omega \cdot \sin\phi \quad (2)$$

$$\sin\beta \cdot \sin\gamma = \cos\omega \quad (3)$$

Thus we have three simultaneous equations from which we can deduce $\alpha\beta\gamma$ on the cross rotations from the ω and ϕ values produced from the self rotation. A program was written which calculated all possible values of α and γ for each β section of the Eulerian map. Table 5.4 shows the results from these calculations using ω, ϕ values of 15, 90 and 105, 90. The first column shows the nearest $\alpha\beta\gamma$ on the output map from ALMN (after unit cell translations if required) and the subsequent columns show the value of the rotation function at that orientation for the monomer, dimer and both datashells.

Table 5.4 Value of the Rotation Function (RF) obtained at points corresponding to (i) $\omega, \phi, \chi = 105, 90, 180$ (or Laue group equivalents) and (ii) $\omega, \phi, \chi = 15, 90, 180$ (or Laue group equivalents)

α	β	γ	Monomer		Dimer		
			RF _{4.5-6}	RF ₆₋₁₀	RF _{4.5-6}	RF ₆₋₁₀	
(1)							
90	15	270	-35	14	-4	19	
45	20	310	-26	-5	-4	-6	
15	20	230	-30	-8	7	-71	
15	20	50	44	36	-4	-6	
45	20	130	3	-28	7	-71	
35	25	320	-19	-12	14	1	
25	25	220	-32	-27	-13	-95	
25	25	40	-14	41	14	1	
35	25	140	-5	-53	-13	-95	
30	30	330	-8	-22	16	1	
30	30	210	4	-35	23	-56	
30	30	30	3	41	16	1	
30	30	150	-19	-21	23	-56	
20	35	335	5	-38	-22	-35	
35	35	205	19	-3	21	-24	
20	35	155	-6	10	29	-24	
40	35	25	41	27	-22	-35	
20	40	335	13	-26	0	-28	
40	40	205	6	8	1	-31	
40	40	25	30	30	0	-28	
20	40	155	-39	6	1	-31	
15	45	340	26	-5	46	16	
45	45	200	5	-22	-9	-58	
45	45	20	38	47	46	16	
15	45	160	-31	-15	-9	-58	
15	50	340	3	-16	39	49	
45	50	200	7	-46	-16	-61	
45	50	20	47 (54)	63	39	49	
15	50	160	-52	-20	-16	-61	
10	55	340	9	-6	-1	67	
50	55	200	5	-36	-12	-35	
50	55	20	8 (67)	51	-1	67	
10	55	160	-38	3	-12	-35	
10	60	345	9	-1	5 (44)	100	
50	60	195	-24	-14	-11	-35	
50	60	15	44	104	5 (44)	100	
10	60	165	5	-14	-11	-35	
5	65	345	40	7	-7	96	
55	65	195	-10	9	6	-33	
55	65	15	-1	99	-7	96	
5	65	165	24	-15	6	-33	
5	70	345	46	15	-4	80	
55	70	195	22	5	-8	-10	
55	70	15	-8	71	-4	80	
5	70	165	-9	5	-8	-10	
5	75	345	19	23	-27	85	
55	75	195	11	12	-1	8	
55	75	15	-7	62	-27	85	
5	75	165	0	0	-1	8	
5	80	345	-9	23	-30	110	
55	80	195	-23	25	-4	32	
55	80	15	-10	73	-30	101	
5	80	165	19	-9	-4	32	
0	85	345	4	13	-12	123	
60	85	195	-11	31	37	53	
60	85	15	-16	101	-12	123	
0	85	165	16	-2	37	53	

† dimer related to (50,60,10)

(ii)						
30	75	90	-8	29	21	-24
30	75	270	36 (82)	-72	21	-24
80	80	80	2	12	41	-30
100	80	100	-28	12	-46	-23
100	80	280	3	-55	41	-30
80	80	260	-7	-8	-46	-23
100	85	75	8	-26	0	-19
80	85	105	5	-47	11	-100
80	85	285	43	-20	0	-19
100	85	255	2	-48	11	-100

None of the peaks predicted for $\omega, \phi = 15, 90$ were very hopeful, all having at least one negative value for the rotation function. However, when one used $\omega, \phi = 105, 90$ one peak in particular looked very promising, it was at $\alpha\beta\gamma = 50, 60, 15$. The other two peaks marked by asterisks are part of the same feature which extends over a large area of Eulerian space. The peak can be seen on the contoured map given in figure 5.10.

A similar procedure was carried out for the smaller self-rotation feature at $\alpha\beta\gamma$ (self) = 30, 20, 30 ($\omega, \phi = 10, 30$ or 100, 30). The results are given in table 5.5. None of the predicted peaks had values of R (rotation function,) significantly above noise level for all four runs.

If we compare the space groups of the monomer and dimer cross rotation functions (see figures 5.8 and 5.9), we see that the dimer two-fold axis creates a symmetry equivalent at $60 - \alpha, \beta, -\gamma$. For the correct $\alpha\beta\gamma$ on the monomer cross rotations we would therefore also expect to see a peak at this position. The predicted dimer equivalent for $\alpha\beta\gamma = 50, 60, 15$ is $\alpha\beta\gamma = 10, 60, 345$ and we can see from table 5.4 that there is not the peak we would expect at this position.

At about this time we were offered some new coordinates from Oxford which had undergone refinement. The R-factor for the refinement was 0.418 using terms to 2.6\AA . These coordinates were used in a similar way to the first set and

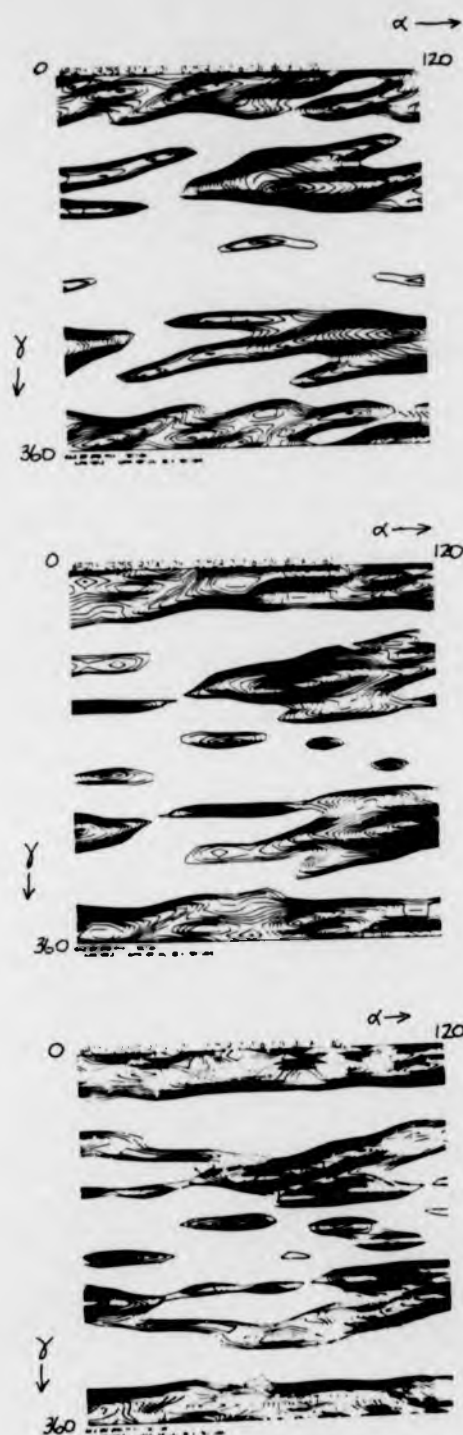


Figure 5.10 Contour Maps of the Cross Rotation Function
Using the 6-10 Å Datashell. Contours every 2.0 above 5.0

Table 5.5 Values of the Rotation Function (RF) obtained at points corresponding to (i) $\omega\phi\chi = 100, 30, 180$ (or Laue group equivalents) and (ii) $\omega\phi\chi = 10, 30, 180$ (or Laue group equivalents).

(i)	α	β	γ	Monomer		Dimer	
				RF _{4.5-6}	RF ₆₋₁₀	RF _{4.5-6}	RF ₆₋₁₀
	30	10	270	-52	54	-45	-17
	100	15	320	-63	68	-20	2
	80	15	220	-83	85	13	13
	80	15	40	44	45	-20	2
	100	15	140	71	13	13	13
	90	20	330	-14	67	25	19
	90	20	210	-9	74	38	5
	90	20	30	27	67	25	19
	90	20	150	71	12	38	5
	80	25	335	15	41	25	12
	100	25	205	15	52	-27	-7
	100	25	25	21	90	25	12
	80	25	155	-19	8	-27	-7
	80	30	340	5	32	-32	15
	100	30	200	-16	15	-37	-23
	100	30	20	-15	92	-32	15
	80	30	160	-17	32	-37	-23
	75	35	340	14	29	-16	27
	105	35	200	16	3	-15	-33
	105	35	20	-13	108	-16	27
	75	35	160	-4	15	-15	-33
	70	40	345	-30	33	-28	29
	110	40	195	-18	-7	0	-34
	110	40	15	-5	101	-28	29
	70	40	165	-36	14	0	-34
	70	45	345	-14	28	1	25
	110	45	195	4	-2	19	-15
	110	45	15	-1	80	1	25
	70	45	165	-14	39	19	-15
	70	50	345	-1	20	5	29
	110	50	195	-24	0	22	0
	110	50	15	-16	71	5	29
	70	50	165	7	68	22	0
	65	55	350	-6	31	-24	40
	115	55	190	-48	-10	-25	-34
	115	55	10	16	84	-24	40
	65	55	170	-16	49	-25	-34
	65	60	350	-73	41	-27	57
	115	60	190	-41	-15	-2	-36
	115	60	10	19	87	-27	57
	65	60	170	10	35	-2	-36
	65	65	350	-45	49	23	69
	115	65	190	-18	-21	5	-20
	115	65	10	27	87	23	69
	65	65	170	-4	21	5	-20
	65	70	350	-1	56	27	82
	115	70	190	-1	-16	-8	8
	115	70	10	3	90	27	82
	65	70	170	4	29	-8	8
	65	75	350	-23	67	9	101
	115	75	190	-17	-3	0	34
	115	75	10	-32	94	9	101
	65	75	170	-16	47	0	34
	60	80	350	-37	56	-7	116
	0	80	190	-32	16	-1	50
	0	80	10	-16	104	-7	116
	60	80	170	-39	62	-1	50
	60	85	350	-47	64	4	110
	0	85	190	-37	33	20	61
	0	85	10	-40	89	4	110
	60	85	170	-1	62	20	61
(ii)							
	90	80	90	-25	-10	3	-44
	90	80	270	16	-18	3	-44
	30	85	80	-11	9	19	-54
	30	85	100	-38	-23	-18	-51
	30	85	280	4	-53	19	-54
	30	85	260	20	-52	-18	-51

the results obtained using them are now discussed.

5.4.1 Cross rotations using new coordinates

The monomer cross rotation for the 4.5-6⁰Å datashell produced 19 peaks with peak heights ranging from 95 to 121 (relative to an origin of 50). The 6-10⁰Å datashell produced 22 peaks with peak heights ranging from 89 to 186. A Patterson sphere of 3-30⁰Å was used for both. Both sets of data were sharpened using a B-factor of -23.4. Table 5.6 lists all of the monomer cross rotation peaks obtained using the new coordinates.

The dimer cross rotation for the 4.5-6⁰Å datashell produced 29 peaks with peak heights ranging from 100 to 139. The 6-10⁰Å datashell produced 16 peaks with peak heights ranging from 105 to 262. The same Patterson sphere of integration and sharpening factor were used for both cross rotation runs. Table 5.7 lists all the dimer cross rotation peaks obtained using the new coordinates. The only peak which showed any consistency between the monomer cross and dimer cross rotations was at approximately $\alpha\beta\gamma = 20, 30, 180$. The spherical polars for this peak and its symmetry equivalents are:

Table 5.6 Cross-rotation peaks obtained using structure factors from the new coordinates for the sheep liver monomer in a P1 cell and the observed bacterial structure factor amplitudes

α	β	γ	RF _{4.5-6}	RF ₆₋₁₀	
25	20	105	97	25	
25	20	160	98	3	
0	25	135	100	18	
15	30	190	110	42	*
40	30	100	121	-11	
105	30	5	-13	115	
75	30	20	-29	95	
20	30	175	54 (89)	102	*
55	30	335	63	103	*
25	35	145	112	-42	
10	35	80	95	2	
55	35	80	110	-16	
100	35	170	98	23	
85	45	70	97	13	
110	45	100	103	29	*
115	45	5	-9	131	
25	45	15	-21	89	
65	45	0	-28	101	
95	50	165	112	86	*
85	50	225	104	-100	
80	50	15	64	122	*
20	50	95	65	97	*
50	55	115	-31	122	
75	55	105	-30	118	
90	55	310	119	-97	
55	60	270	20	93	
60	65	10	8	110	
90	65	255	28	133	
15	65	45	102	11	
55	65	155	101	12	
115	70	310	120	-84	
10	70	105	11	186	
0	75	10	-27	117	
55	75	265	56	98	*
5	80	50	96	-4	
20	80	330	114	-39	
10	85	210	-85	139	
65	85	350	-22	115	
75	85	125	-36	101	
40	90	155	1	122	
35	90	215	6	130	

Table 5.7 Cross-rotation peaks obtained using structure factors from the new coordinates for the sheep liver dimer in a P1 cell and the observed bacterial structure factor amplitudes

α	β	γ	RF _{4.5-6}	RF ₆₋₁₀	
105	10	115	101	36	
70	15	180	-21	166	
5	20	155	112	51	
15	20	200	111	-6	
20	25	105	114	65	
30	25	165	108	41	
70	30	135	104	-100	
20	30	180	104	126	* +
20	35	120	110	-17	
95	35	165	117	-45	
65	35	190	139	31	
15	40	20	119	127	*
5	40	85	120	82	*
15	40	180	-58	135	
70	45	0	11	242	
110	45	10	-16	262	
110	45	95	48	129	
50	45	175	17	141	
90	50	0	86	258	*
25	50	105	-69	105	
65	50	165	15	137	
45	50	165	102	40	
90	50	165	101	6	
75	55	55	127	-49	
70	55	75	132	22	
45	60	40	104	-100	
90	60	50	114	-100	
30	60	55	123	-35	
75	60	0	-39	183	
85	65	105	16	200	
5	65	55	109	-33	
20	70	65	108	-100	
40	75	50	111	50	
115	75	30	107	21	
65	75	45	127	-72	
90	75	180	101	196	*
70	75	245	38	228	
100	80	50	100	-100	
20	80	45	109	40	
50	80	55	111	21	
10	80	95	102	79	
65	85	350	22	240	
55	90	45	0	168	
25	90	205	36	196	

ω	ϕ	χ
169.7	-165.0	150.5
145.7	-105.0	35.9
14.0	- 45.0	91.7
82.9	135.0	165.0
99.7	15.0	174.8
92.6	75.0	160.7

None of these have the expected $\omega\phi = (105, 90)$ (or $(15,90)$, $(10,30)$, $(100,30)$).

By comparing matrix elements for the rotation matrix we can obtain an expression relating the directional cosines to the Eulerian angles, namely:

$$\cos\beta = \cos\chi (1 - DC3^2) + DC3^2 \quad (4)$$

$$\sin\alpha = \frac{DC2 \cdot DC3 (1 - \cos\chi) - DC1 (\sin\chi)}{\sin\beta} \quad (5)$$

$$\sin\gamma = \frac{DC2 \cdot DC3 (1 - \cos\chi) + DC1 (\sin\chi)}{\sin\beta} \quad (6)$$

By substituting the directional cosines and a chi value of 180° into these expressions we were able to calculate the $\alpha\beta\gamma$ on the self rotation map corresponding to any orientation given by the output peaks on the cross rotation maps. Using these expressions we found that the position on the self rotation map corresponding to the $\alpha\beta\gamma = 20, 30, 180$ cross peak was $\alpha\beta\gamma$ (self) = 74.98, 28, 74.98. Figure 5.11 shows the position of this compared to the actual peak on

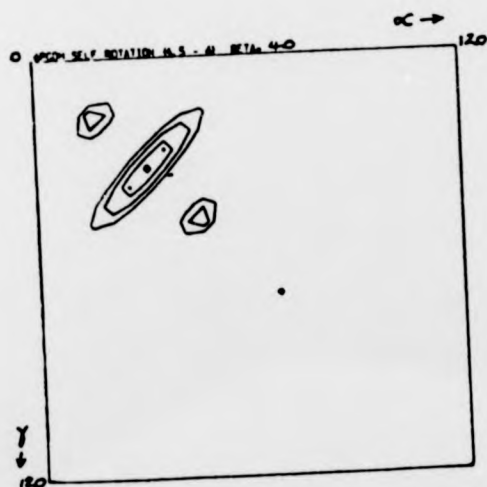
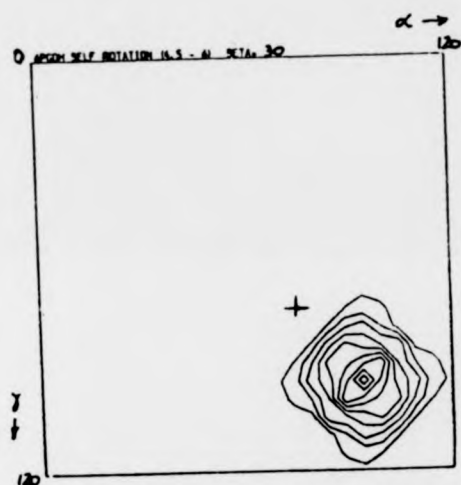
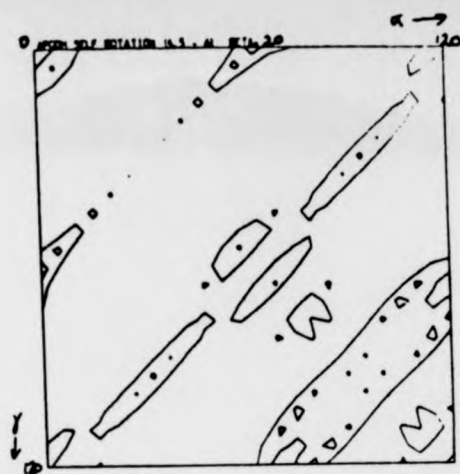


Figure 5.11 Contour Map of the Self Rotation Function
Using the 4.5-6 Å Datashell Showing the $\alpha\beta\gamma = 74.98, 28, 74.98$
Position Relative to the $\alpha\beta\gamma = 90, 30, 90$ Peak

the self rotation map.

Equations (4), (5) and (6) were also used on the three most hopeful of the major cross peaks (those marked by + in table 5.2) using the old coordinates and all their symmetry equivalents). The results are given in table 5.8. None of the predicted positions were very close to the feature at $\alpha\beta\gamma$ (self) = 90, 30, 90.

The peak at $\alpha\beta\gamma$ (cross) = 20, 30, 180 did not feature in the old coordinate runs but was reasonably high (98) for the dimer 6-10 $\overset{\circ}{\text{\AA}}$ run. The $\alpha\beta\gamma$ (cross) = 50, 60, 15 peak which had large values of R for the monomer, dimer and both datashells using the old coordinates was also present for both the monomer and dimer runs using the 6-10 $\overset{\circ}{\text{\AA}}$ datashell from the new coordinates. However, it was not present when the 4.5-6 $\overset{\circ}{\text{\AA}}$ datashell from the new coordinates were used, or at the dimer related position on the monomer runs (old coordinates).

5.4.2 Summary

The cross rotation functions were far more susceptible to false peaks than the self rotation function. There were large variations in peak heights and position when one changed the calculating parameters even slightly. For this reason the main criterion used in evaluating the cross rotation peaks was the ability to survive such changes.

Table 5.8 The 'best' old coordinate peaks from table 5.2 and their corresponding values of $\alpha\beta\gamma(\text{self})$.

CROSS						SELF		
α	β	γ	DC1	DC2	DC3	α	β	γ
35	45	345	-0.3898	0.8358	0.3866	65.0	134.51	65.0
155	45	345	-0.4018	0.0352	0.9150	5.0	47.6	5.0
95	-45	165	0.2728	0.3896	-0.8796	-55.0	56.81	-55.0
85	135	165	-0.6087	-0.7254	-0.3213	50.0	142.52	50.0
25	-135	345	0.3418	-0.9391	0.0361	-70.0	175.86	-70.0
145	-135	345	0.9220	-0.1626	0.3515	-10.0	138.84	-10.0
55	60	30	-0.1406	0.6343	0.7602	77.5	81.04	77.5
175	60	30	0.4855	-0.1531	-0.8608	17.51	61.19	17.51
115	-60	-150	0.6539	0.5992	-0.4619	-42.5	124.98	-42.5
65	120	-150	-0.8885	-0.2801	-0.3634	17.5	137.38	17.5
5	-120	30	-0.2132	-0.9619	0.1711	-77.5	160.3	-77.5
125	-120	30	0.6423	-0.5885	0.4910	-42.5	121.19	-42.5
110	80	125	-0.0897	-0.6813	-0.7265	82.51	86.81	82.51
50	-80	-55	0.7923	-0.6079	-0.0519	37.49	174.05	37.49
170	-80	-55	0.6516	0.2699	0.7089	22.5	89.71	22.5
10	100	-55	-0.5116	0.8030	-0.3057	-57.5	144.4	-57.5
130	100	-55	-0.8897	-0.0388	0.4549	-2.5	125.84	-2.5
70	-100	125	0.3550	0.6819	-0.6395	-62.5	100.49	-62.5

The only peak which corresponded to the self rotation result and showed the ability to survive change in the calculating parameters and datashells used was at $\alpha\beta\gamma$ (cross) = 50, 60, 15. This peak has the spherical polars $\omega\phi\chi$ = 75, -90, 62.4 (\equiv 105, 90, -62.4). This implies that in the bacterial P3₂21 cell the dimer axis is at an angle of 105° to c, 90° to a and 30° to b and that to orientate the molecules with respect to the sheep liver atomic coordinates a further rotation of -62.4° about that axis is required.

This result is at odds with our earlier comment that the dimer axis was likely to be orientated towards the longest cell axis (see section 5.3). However, the rotation function result is stronger evidence than our earlier speculation. It must be coincidental that the longest cell axis in the two different forms are approximately the same length. Dodson et al (1966) found that in rhombohedral insulin the non-crystallographic 2 fold axes also arranged themselves perpendicular to the crystallographic 3-fold axis.

One of the main problems with using Crowther's FFT program is the need to use a spherical volume of integration. This is necessary because the program expands the Patterson density in terms of spherical harmonics rather than Cartesian Fourier components. As we have already stated the sheep liver 6PGDH molecule is ellipsoidal and therefore any sphere of integration that we choose will

either exclude many of the vectors within the molecule, or include a high proportion of intermolecular vectors which will not be matched by the non-crystallographic rotation. Either of these significantly increase the noise level. This is the most probable cause of error in our cross rotation runs and probably explains the lack of a peak at the dimer related position.

Any structural differences between the two molecules will obviously reduce the power of the rotation function. There is certain to be some structural differences between the enzyme extracted from a thermophilic bacterium and from sheeps' liver. Thus the cross rotations will be more prone to error than the self rotations. It is of course hoped that the two structures are similar enough to allow the molecular replacement technique to work. The persistence of the $\alpha\beta\gamma$ (cross) = 50, 60, 15 peak is evidence to support this hope. Cutfield et al (1974) successfully used the cross rotation technique when solving the Insulin structure from Hagfish (*Myxine glutinosa*) by rotation of the known 2-zinc pig insulin F_C^2 onto the Hagfish F_O^2 . The amino acid sequence (Emdin et al (1973)) indicates changes in more than a third of the amino acids when the two forms of insulin are compared.

CHAPTER 6

THE TRANSLATION PROBLEM

6.1 Introduction

The results of applying the Rotation Function to our data allowed us to define the best orientation of the bacterial electron density with respect to the cell axes, using the sheep liver enzyme structure as a model (see chapter five). To fully position our model in the bacterial cell, however, we are also required to know the translational parameters. There are two main ways of determining these vectors; the first (earlier) method is to use the scalar product of the Patterson cross vector peaks as described by Rossmann et al (1964). This method suffers from two main faults; firstly one must successfully separate the inter- and intra-molecular vectors in the Patterson which, as we have already pointed out is difficult for an elliptical molecule. Secondly, the component of the translation vector derived is always poorly defined perpendicular to the non crystallographic rotation axis. This is clearly demonstrated in the results published using this method by Dodson et al (1966) during the study of Rhombohedral Zinc Insulin. The second method for obtaining the translational vector is to conduct an R-factor search

over the unit cell for an oriented molecule. The merits of both of these methods have been investigated by Nixon (1973). He has shown that the R-factor search method is less noisy and more precise than the Patterson scalar product method (Nixon and North (1976)). The method has been successfully used for structure determinations by Cutfield et al (1974) and Ru-chang et al (1983).

6.2 Method

The translation parameters were searched for by looking for a minimum R-value between the observed amplitudes and the different sets of calculated structure factor amplitudes generated as the oriented molecule was translated through the unit cell. The R-factor calculated was defined by the following expression:-

$$R = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$$

The program used to perform the R-factor search was written by E.Dodson (pers. comm.) and is called SEARCH. The program uses the fact that if the partial calculated structure factors ($F_{C1}, F_{C2}, F_{C3}, \dots, F_{CNSYM}$) are obtained for all symmetry equivalent positions from the model, then the value of the overall calculated structure factor F_{CALC} for any set of symmetry related translations ($t_1, t_2,$

$\underline{t}_3 \dots \underline{t}_{\text{NSYM}}$) is given by:-

$$\underline{F}_{\text{CALC}}(\text{hkl}) = \underline{F}_{\text{C1}}(\text{hkl}) \exp(-2\pi i \underline{h} \cdot \underline{t}_1) + \underline{F}_{\text{C2}}(\text{hkl}) \exp(-2\pi i \underline{h} \cdot \underline{t}_2) + \dots + \underline{F}_{\text{CNSYM}}(\text{hkl}) \exp(-2\pi i \underline{h} \cdot \underline{t}_{\text{NSYM}})$$

Therefore once the partial $\underline{F}_{\text{Ci}}$ values are obtained it is only necessary to sum them together with appropriate phase modifications to generate the different sets of $\underline{F}_{\text{CALC}}$'s required for the R-factor calculation.

The sheep liver atomic coordinates were rotated using the rotation matrix obtained from ALMN (see chapter five), by means of the program LSQAB. These rotated coordinates were then used to calculate a set of structure factors for a single dimer in a P1 cell. The program CAD was then used to produce 6 sets of partial $\underline{F}_{\text{Ci}}$'s for the bacterial cell from these data. The partial $\underline{F}_{\text{Ci}}$'s were then input into SEARCH and an R-factor search undertaken. It was necessary to calculate $\underline{F}_{\text{Ci}}$'s for both a P3₁21 cell and a P3₂21 cell as we could not at this stage determine which one of this enantiomorphic pair the bacterial crystals had formed in. Because the Rotation Function calculations were performed in Patterson space this problem did not arise earlier as both space groups have the same Laue group ($\bar{3}m$).

6.3 Results

It was found that the output R-factor maps displayed x,

y, $1/2 + z$ symmetry and that it was therefore only necessary to search over half a unit cell. The initial searches were made on a 1\AA grid using data to 4.5\AA . This required over a million R-factor calculations to be made for each space group, each calculation based on data from over 4,000 reflections. Individual runs of the program were done for each z-section and tables 6.1 and 6.2 summarise the results. The first column gives the grid points for the lowest R-factor on the z-section (given approximately in \AA), the second column gives the value of R in percent and the third column shows the average R value for all 14641 grid points on that section.

6.3.1 Results for the P3₁21 cell

The average R-factor for each z-section was $62.51 \pm 0.58\%$. There were four points out of the 1,062,075 points inspected which had R-values below 54.0%, these are marked with asterisks in table 6.1. A translation search on a 0.2\AA grid was conducted around each of these four points and the results are shown in table 6.3. The lowest R-value obtained was 52.75%, this was for a translation of 4.31\AA in x, 64.63\AA in y and 25.40\AA in z. No one minimum was strikingly low compared to the others and none of them were of the order of 45% that one would expect for a good correlation between structure and model (E.Dodson, pers. comm.).

Table 6.1 Results from the translation SEARCH using a P3₁21⁰ cell. The grid points are approximately Angstrom units.

GRID POINTS			R MIN	R SECTION
80	1	0	53.48	62.45
35	71	1	54.23	62.56
35	22	2	54.25	62.51
9	7	3	57.66	61.93
89	88	4	54.31	62.54
44	115	5	55.20	62.62
66	80	6	55.23	62.55
114	93	7	55.76	62.49
85	43	8	54.88	62.48
104	44	9	55.79	62.59
65	62	10	55.60	62.48
13	114	11	54.69	62.23
95	0	12	54.76	62.53
2	116	13	54.84	62.49
8	96	14	55.00	62.48
30	98	15	54.99	62.39
17	72	16	55.48	62.51
117	52	17	55.19	62.75
65	62	18	54.73	62.47
115	112	19	54.22	62.41
25	18	20	53.94	62.52
31	4	21	55.27	62.68
116	116	22	54.27	62.63
72	58	23	54.25	62.36
16	98	24	56.28	62.62
35	79	25	56.00	62.61
4	63	26	53.92	62.36
119	105	27	54.90	62.43
103	9	28	54.82	62.35
34	2	29	55.43	62.50
32	4	30	56.09	62.55
14	68	31	54.14	62.53
23	72	32	55.72	62.75
22	53	33	54.77	62.71
79	75	34	55.63	62.59
75	32	35	54.94	62.63
65	93	36	55.04	62.36
75	32	37	55.26	62.46
74	93	38	55.88	62.41
120	82	39	54.53	62.33
112	114	40	54.90	62.47
15	86	41	54.70	62.52
74	25	42	54.87	62.68
51	117	43	55.44	62.53
48	105	44	55.07	62.51

continued

Table 6.1 continued

29	81	45	55.37	62.56
30	59	46	55.09	62.60
83	102	47	55.19	62.70
49	70	48	55.84	62.63
111	109	49	54.46	62.70
39	50	50	54.92	62.64
38	49	51	53.61	62.61
103	10	52	55.92	62.59
11	27	53	55.55	62.34
83	11	54	54.15	62.34
37	62	55	55.32	62.46
16	44	56	55.43	62.56
20	43	57	55.60	62.53
44	52	58	55.02	62.48
38	23	59	54.72	62.54
34	116	60	55.61	62.65
22	12	61	56.03	62.41
56	58	62	54.64	62.35
103	11	63	55.88	62.48
45	101	64	54.63	62.53
25	87	65	54.87	62.38
4	108	66	55.05	62.39
29	62	67	55.80	62.64
74	62	68	55.50	62.59
45	84	69	55.36	62.41
98	28	70	55.44	62.31
90	76	71	55.31	62.61
75	63	72	54.53	62.66
115	63	73	56.12	62.41
96	1	74	54.96	62.47
80	1	75	53.47	62.45

Table 6.2 Results from the translation SEARCH using a $P_{3,21}$ cell. The grid points are approximately in \AA units.

GRID POINTS \AA			R MIN	R SECTION
2	2	0	53.98	61.26
79	14	1	54.14	61.44
78	39	2	53.90	61.54
101	66	3	52.61	61.43
26	30	4	53.42	61.69
106	101	5	54.78	61.83
105	87	6	54.06	61.53
109	27	7	53.65	61.13
40	9	8	54.21	61.30
60	89	9	53.05	61.34
116	74	10	53.62	61.36
79	89	11	53.50	61.38
16	48	12	54.37	61.63
43	48	13	54.55	61.79
7	79	14	54.64	61.65
98	8	15	54.09	61.23
32	37	16	52.99	61.23
46	110	17	53.97	61.30
70	75	18	53.83	61.34
86	40	19	52.94	61.24
53	22	20	54.44	61.61
78	89	21	53.88	61.61
6	74	22	53.71	61.39
92	6	23	53.82	61.08
40	16	24	54.04	61.03
9	106	25	54.16	61.14
63	99	26	53.99	61.12
11	83	27	54.50	61.02
10	63	28	54.09	61.29
47	110	29	54.27	61.74
47	110	30	54.30	61.65
83	83	31	53.68	61.52
51	74	32	53.87	61.44
21	25	33	54.21	61.55
33	49	34	54.39	61.50
102	54	35	52.55	61.05
45	89	36	53.80	61.29
61	105	37	53.81	61.64
79	89	38	53.85	61.38
101	65	39	52.61	61.26
101	65	40	54.21	61.23
98	79	41	54.36	61.37
56	8	42	52.98	61.18
53	7	43	53.85	61.19

continued

Table 6.2 continued

28	23	44	53.67	61.65
58	49	45	55.08	61.81
86	70	46	53.46	61.57
61	106	47	52.61	61.52
34	21	48	53.36	61.40
47	110	49	53.84	61.47
6	120	50	52.68	61.39
102	81	51	54.50	61.23
95	29	52	54.06	61.35
46	110	53	53.57	61.24
1	88	54	54.54	61.45
101	66	55	51.79	61.17
103	63	56	53.50	61.17
79	89	57	54.79	61.36
56	8	58	53.94	61.15
61	93	59	53.36	61.31
45	6	60	53.17	61.54
85	74	61	53.39	61.45
84	48	62	53.34	61.59
101	65	63	53.13	61.33
10	55	64	53.34	61.37
80	112	65	53.91	61.15
5	26	66	53.42	61.01
85	49	67	52.25	61.03
32	49	68	52.79	61.22
95	111	69	53.91	61.46
18	88	70	51.17	61.61
100	16	71	54.02	61.20
25	47	72	53.37	61.40
78	89	73	54.92	61.27
3	68	74	54.12	60.79
2	2	75	53.98	61.26

Table 6.3 Results of a fine grid translational SEARCH around the positions of the minima obtained using a $P3_121$ cell (Table 6.1). The grid used was approximately 0.2 Å.

ORIGINAL GRID POINT	R(1Å)			R(0.2Å)			TRANSLATION IN Å		
	x	y	z	x	y	z	x	y	z
80 1 0	53.48	53.46	82.27	1.03	0.00				
25 18 25	53.94	53.04	25.85	18.88	19.69				
4 63 26	53.92	52.75	4.31	64.63	25.40				
38 49 51	53.61	53.14	38.98	50.06	50.02				

Table 6.4 Summary of results from the program PRJANG using a $P3_121$ cell

Translation	Symmetry Element	Number of close contacts
82.27 1.03 0.00	(1) x,y, z	0
	(2) -y,x-y,z+1/3	64
	(3) y-x,-x,z+2/3	64
	(4) y,x,-z	0
	(5) x-y,-y,2/3-z	20
	(6) -x,y-x,1/3-z	112
	TOTAL	260
25.85 18.88 19.69	(1)	0
	(2)	127
	(3) AS ABOVE	127
	(4)	394
	(5)	0
	(6)	90
	TOTAL	738
4.31 64.63 25.40	(1)	0
	(2)	43
	(3) AS ABOVE	43
	(4)	0
	(5)	186
	(6)	104
	TOTAL	376
38.98 50.06 50.02	(1)	0
	(2)	20
	(3) AS ABOVE	20
	(4)	55
	(5)	903
	(6)	0
	TOTAL	998

Table 6.3 Results of a finegrid translational SEARCH around the positions of the minima obtained using a $P3_121$ cell (Table 6.1). The grid used was approximately 0.2 Å.

ORIGINAL GRID POINT	R(1Å)			R(0.2Å)			TRANSLATION IN Å		
	x	y	z	x	y	z	x	y	z
80 1 0	53.48	53.46	82.27	1.03	0.00				
25 18 25	53.94	53.04	25.85	18.88	19.69				
4 63 26	53.92	52.75	4.31	64.63	25.40				
38 49 51	53.61	53.14	38.98	50.06	50.02				

Table 6.4 Summary of results from the program PRJANG using a $P3_121$ cell

Translation	Symmetry Element	Number of close contacts
82.27 1.03 0.00	(1) x,y, z	0
	(2) -y, x-y, z+1/3	64
	(3) y-x, -x, z+2/3	64
	(4) y, x, -z	0
	(5) x-y, -y, 2/3-z	20
	(6) -x, y-x, 1/3-z	112
	TOTAL	260
25.85 18.88 19.69	(1)	0
	(2)	127
	(3) AS ABOVE	127
	(4)	394
	(5)	0
	(6)	90
	TOTAL	738
4.31 64.63 25.40	(1)	0
	(2)	43
	(3) AS ABOVE	43
	(4)	0
	(5)	186
	(6)	104
	TOTAL	376
38.98 50.06 50.02	(1)	0
	(2)	20
	(3) AS ABOVE	20
	(4)	55
	(5)	903
	(6)	0
	TOTAL	998

One method of checking the results of a translation function is to apply the translation vector obtained, in conjunction with the correct rotation matrix, to the atomic coordinates of the known structure (model) and then check for overlap when these transformed coordinates are placed in the observed cell together with the appropriate symmetry. There is a program called PRJANG which checks for such overlaps. Eleanor Dodson kindly ran this program for each of the four translations listed in table 6.3. The program used only the C_a atomic coordinates and flagged all close contact that were less than 5\AA . The results are summarised in table 6.4.

We will expect some close contacts to occur when we place the model structure into the bacterial cell and apply the appropriate symmetry (unless the model is perfect). It is difficult to quantify how much overlap one can deem as acceptable in these circumstances. Ideally each of the likely translations should be applied separately and the resulting structures closely examined, especially in the regions of close contacts. The use of an Evans and Sutherland PS300 interactive graphics computer would greatly speed up this process if this approach could be justified. In any case, without we can only draw qualitative conclusions of the results for the R-factor search by saying that translations (82.27, 1.03, 0.0) and (4.31, 64.63, 25.40) seem far more likely than (38.98, 50.06, 50.02) and (25.85, 18.88, 19.69). We can also make a qualitative

judgement about which of the two possible space groups is the correct one (see section 6.3.2).

6.3.2 Results for the $P3_221$ cell

The average R-factor for each z-section was $61.37 \pm 0.46\%$. The seven points with the lowest R-factors all had values of 52.55% or less, these are marked with asterisks in table 6.2. A 0.2\AA grid translation search was conducted around each of these seven points and the results are shown in table 6.5. The lowest R-factor obtained was 50.30% for a translation of 104.02\AA in x, 67.50\AA in y and 2.95\AA in z. Again no minimum was strikingly low compared to the others or close to 45% (see section 6.3.1). These translations were also input into the program PRJANG to check for close contacts and the results are summarised in table 6.6.

If one compares the results obtained using a $P3_221$ cell (table 6.6) and a $P3_121$ cell (table 6.4), it is clear that there are far fewer overlaps resulting from the translation-containing symmetry operations (2, 3, 5 and 6 in figures 6.4 and 6.6) for the $P3_221$ cell. This would seem to indicate that the $P3_221$ cell is the correct one.

Table 6.5 Results of a fine grid translational SEARCH around the positions of the minima obtained using a $P3_221$ cell (Table 6.2). The grid used was approximately 0.2 \AA .

ORIGINAL GRID POINT	$R(1\text{\AA})$	$R(0.2\text{\AA})$	TRANSLATION IN \AA		
			x	y	z
101 66 3	52.61	50.30	104.02	67.50	2.95
102 54 35	52.55	51.45	104.22	54.98	34.46
101 65 39	52.61	51.68	103.81	66.68	38.20
61 106 47	52.61	51.91	62.58	108.53	46.28
101 66 55	51.79	51.53	103.61	67.50	54.15
85 49 67	52.25	51.28	86.99	50.06	66.36
18 88 70	51.17	51.08	18.46	90.07	68.92

PRJANG using a P3.21 cell

Translation	Symmetry element	Number of close contacts
104.02 67.50 2.95	(1) x,y,z	0
	(2) -y,x-y,z+2/3	0
	(3) y-x,-x,z+1/3	0
	(4) y,x,-z	290
	(5) x-y,-y,1/3-z	166
	(6) -x,y-x,2/3-z	0
	TOTAL	456
104.22 54.98 34.46	(1)	0
	(2)	0
	(3) AS ABOVE	0
	(4)	0
	(5)	1138
	(6)	0
	TOTAL	1138
103.81 66.68 38.20	(1)	0
	(2)	0
	(3) AS ABOVE	0
	(4)	0
	(5)	585
	(6)	0
	TOTAL	585
62.58 108.53 46.28	(1)	0
	(2)	214
	(3) AS ABOVE	214
	(4)	0
	(5)	8
	(6)	87
	TOTAL	523
103.61 67.50 54.12	(1)	0
	(2)	0
	(3) AS ABOVE	0
	(4)	0
	(5)	179
	(6)	6
	TOTAL	185
86.99 50.06 66.36	(1)	0
	(2)	91
	(3) AS ABOVE	91
	(4)	266
	(5)	154
	(6)	635
	TOTAL	1237
18.46 19.07 68.92	(1)	0
	(2)	225
	(3) AS ABOVE	225
	(4)	0
	(5)	0
	(6)	0
	TOTAL	450

CHAPTER 7CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK7.1 Concluding Remarks

None of the minimum R-values produced by the translation search were markedly lower than the minima on other sections (see tables 6.1 and 6.2). This is probably due to the fact that the results obtained for the cross rotations were never as strong and unambiguous as our self-rotation result. The sensitivity of the cross-rotation peaks to changes in calculating parameters has already been discussed in section 5.4.2. In consequence, it was decided that we could not confidently position the sheep liver atomic coordinates in the bacterial cell and gain a satisfactory model from which to produce an F_{OBS}^a F_{CALC} electron density map.

There are several factors which may have been separately or collectively obstructive in our attempts to use molecular replacement to solve the bacterial 6PGDH structure. The first factor, which we have already mentioned, is that a spherical volume of integration was required to be used in the rotation functions (see section 5.4.2). This means that one must either exclude large

numbers of intramolecular vectors or include many intermolecular vectors in the calculations. This problem is naturally more prevalent in cross rotations.

When the minimum R-search was carried out one had to use a single rotation matrix which could not be allowed to vary. The rotation matrix could not be obtained with a precision of better than a few degrees and any inherent errors in this matrix obviously reduce the power of the translation search.

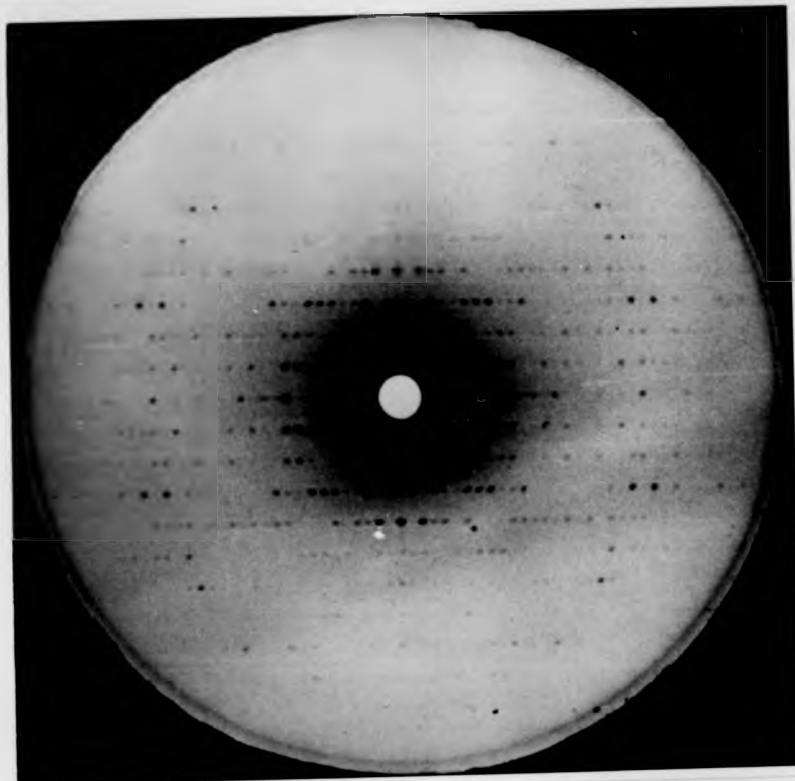
Any structural differences between the bacterial and model structures will naturally reduce the power of molecular replacement techniques. There are likely to be structural differences between the sheep liver and bacterial 6PGDH for two reasons, firstly they are widely separated on the evolutionary scale and secondly there must be some structural features which lead to the enhanced thermal stability in the enzyme extracted from B. stearothermophilus. The changes in structure leading to thermal stability are likely to be subtle (see chapter 1). A large change in thermal stability has been reported as a result of the substitution of a single amino acid residue (tyrosine 88 is replaced by a histidine) in lysozyme extracted from Bacteriophage T4 (Matthews (1976)). In the mutant strain in which the histidine is present the activity at 37°C is about 40% of that of the wild-type enzyme, however at 20°C the activity of the mutant is less than

0.01% of that of the wild type. If we compare the amino acid composition of 6PGDH from B. stearothermophilus and sheep liver (see table 1.2) we can see that they are very similar. We can therefore assume that their overall structures have not drifted too far apart during the course of evolution.

Another possible cause of error is that the bacterial crystals were grown from the holo-enzyme but the atomic coordinates used in the molecular replacement calculations were from the apo-form of the sheep liver enzyme. There are no sheep liver holo-enzyme coordinates available. Co-enzyme binding studies of the sheep liver enzyme at 6\AA resolution (Abdallah et al (1979)) did not show there to be any conformational change when the co-enzyme was bound. Some difficulty has been experienced in binding the co-enzyme without conformational change at 3\AA resolution (Adams (1984) pers. comm.).

In the sheep liver crystals there was sometimes a spontaneous transformation to a different crystal form observed, type A to type B. This transformation between crystal types was characterised by distinct intensity changes which can be illustrated with reference to hol zone precession photographs (see figure 7.1). The type B crystal, which could be induced reversibly by change of salt from sulphate to tartrate formed in the same space group ($C222_1$) but had unit cell dimensions of $a = 71.9\text{\AA}$, $b =$

(A)



(B)

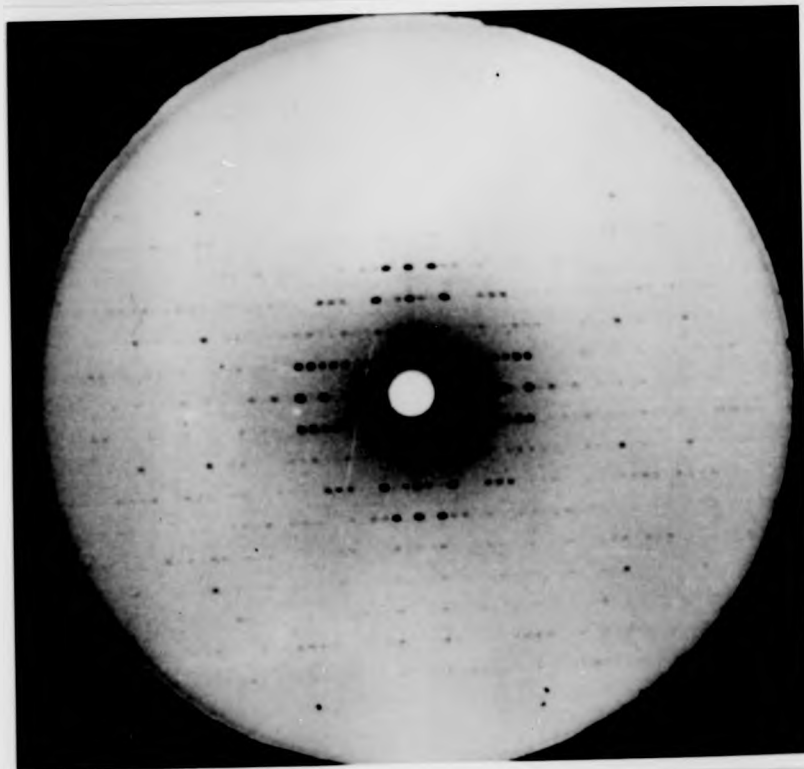


Figure 7.1 16° Precession Photographs of Sheep Liver 6PGDH

(A) Shows Crystal Type A (B) Shows Crystal Type B.

The a^* Axis is Vertical and Attention is Drawn to the Intensity Variation Along This Direction

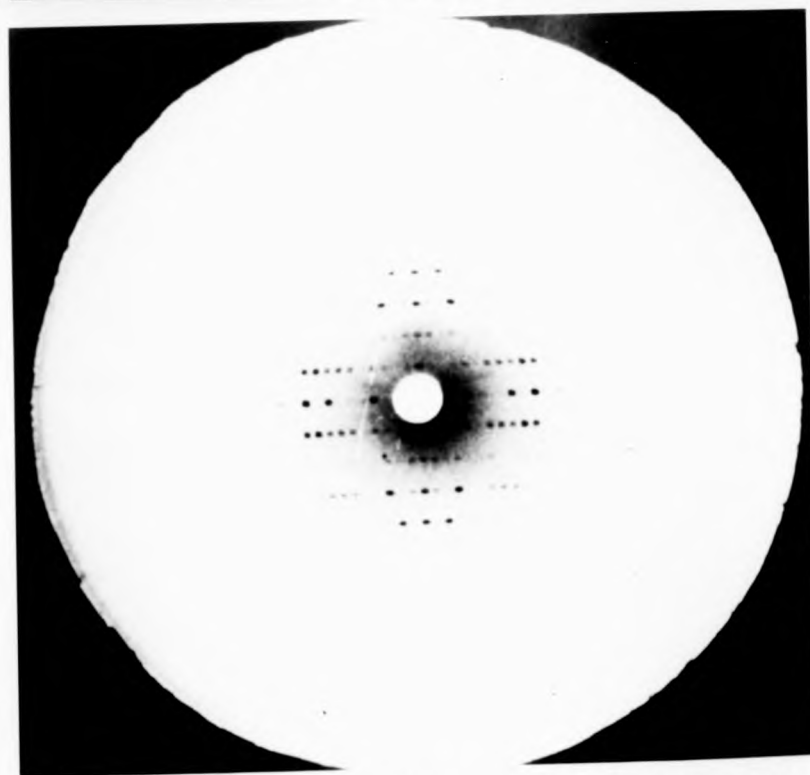
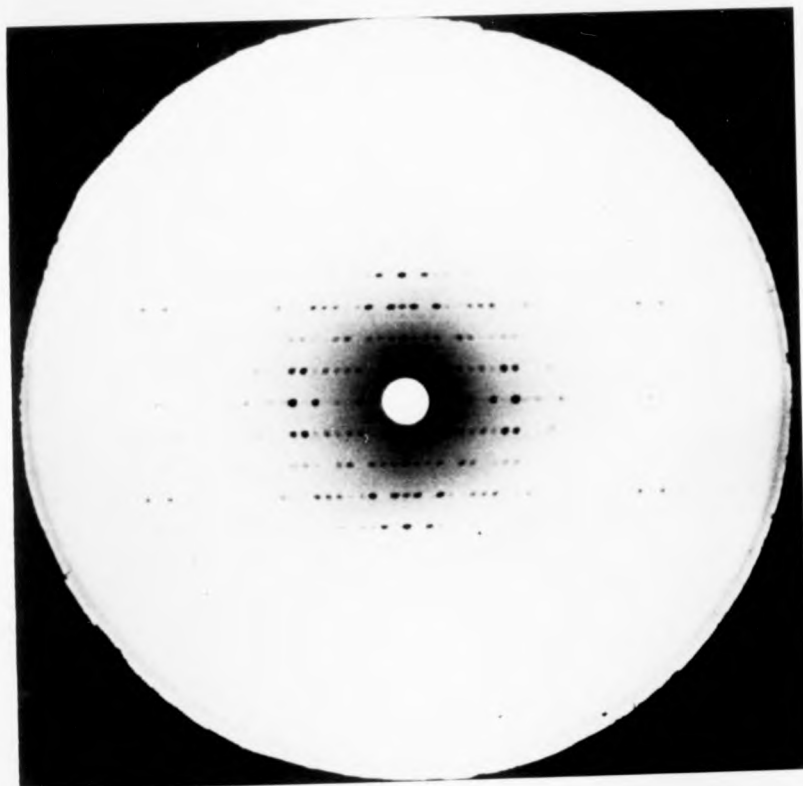


Figure 7.1 16° Precession Photographs of Sheep Liver 6PGDH

(A) Shows Crystal Type A (B) Shows Crystal Type B.

The a^* Axis is Vertical and Attention is Drawn to the Intensity Variation Along This Direction

$146.8\overset{\circ}{\text{\AA}}$, $\underline{c} = 100.6\overset{\circ}{\text{\AA}}$ (c.f. $\underline{a} = 72.72\overset{\circ}{\text{\AA}}$, $\underline{b} = 148.15\overset{\circ}{\text{\AA}}$, $\underline{c} = 102.91\overset{\circ}{\text{\AA}}$). The structure of the B-type crystals has not been determined to high resolution but using $6\overset{\circ}{\text{\AA}}$ data a difference Fourier map indicates tht the phenomenon involves a rotation of the two monomers with respect to each other about an axis parallel to \underline{c} (Helliwell (1977)). If the bacterial crystals naturally occur in the B-form this may lead to problems when one uses A-form coordinates to conduct the cross rotation calculations.

A further cause of conformational change between the sheep liver and bacterial structures of 6PGDH could be due to a 'hinging' motion being allowed about the segment of polypeptide chain joining the large and the small domains of the enzyme. A hinging of this type was observed in the horse liver alcohol dehydrogenase (LADH) structure (Eklund et al (1981)). In the LADH case the two domains undergo a rotation of 7.5° between the apo and holo forms of the enzyme. However, as no conformational change was observed when co-enzyme binding was studied for the sheep liver 6PGDH (Abdalleh et al (1979)) we would not expect simple co-enzyme binding to result in a movement of the subunits about this hinge.

Similar problems have been encountered using molecular replacement techniques on other structures. Smith, Hendrickson and Addison (1983) report that self-rotation calculations allowed the threefold axis in trimeric

haemerythrin to be determined but that cross rotation calculations using myohaemerythrin as a model structure failed to reveal subunit orientations. Cross rotation calculations failed to give convincing peaks in the study of gramicidin (E.Dodson, pers. comm.). A 'post mortem' on the data after the structure had been solved using MIR techniques showed that a region of higher density was in the correct position on the output maps but it could not be made into a convincing peak by simply adjusting the calculating parameters.

7.2 Suggestions for Further Work

Due to the many problems that have beset our attempts to solve the bacterial structure by means of molecular replacement, the first suggestion for further work is that heavy atom derivatives be used to determine the structure by means of multiple isomorphous replacement (MIR). To this end two datasets have now been collected (Helliwell and Habash (1984) pers. comm.). The heavy atom complexes used were $K_2Pt(CN)_4$ and $KAu(CN)_2$. Derivatives obtained with these complexes have successfully been used in the structure determination of sheep liver 6PGDH (Adams et al (1983)). The crystals used were of poor quality and useful data is of a low resolution (approximately 6\AA). If any further crystallographic study of 6PGDH from B. stearothermophilus is undertaken it is essential to grow some fresh crystals.

If a supply of new crystals were available it would also be interesting to see if the use of different salts in the mother liquor such as those reported by Helliwell (1977) will give rise to a transition of crystal types similar to that which occurs in sheep liver 6PGDH (see figure 7.1).

If the problems currently being experienced in binding the co-enzyme to the sheep liver 6PGDH at 3⁰A resolution are found to be due to a hinging of the domains, cross rotation calculations made using individual domains as models may be more successful than the full structure cross rotations that we have conducted. To separate the differences in structure between sheep liver and B. stearothermophilus 6PGDH that arise due to evolution from those which enhance thermal stability, a parallel crystallographic study of the enzyme extracted from the mesophilic bacterium E. coli in conjunction with the sequencing of its amino acids would be useful.

The use of synchrotron radiation was very important in this study because when we tried to collect X-ray data on a conventional source (Elliot GX20) it took an hour to produce setting stills of a usable intensity and the crystals ceased to diffract to high angles after 2-3 hours. This poor ability to diffract was due to two main factors; firstly the crystals were small ($< 0.15\text{mm}$) and had a relatively large unit cell volume ($123.1 \times 123.1 \times 147.7 \text{ \AA}^3$). Secondly they were radiation sensitive (see section 4.2.8). The use

of high flux synchrotron X-ray sources increase the effective crystal lifetime for most proteins by a factor of two to six times (Wilson et al (1983)). It was for this reason that a large amount of my time on this project was spent assisting in the designing and commissioning of the central facility for protein crystallography data collection on beamline seven of the Synchrotron Radiation Source (SRS), Daresbury, UK (Helliwell et al (1982) - see appendix II). It is foreseen that synchrotron radiation will be of importance for any future work on 6PGDH crystals from B. stearothermophilus.

REFERENCES

- Achari A., Heaphy S., Marshall S.E., Muirhead H. (1983), 1982-1983 Daresbury Laboratory Annual Report.
- Adams M.J., Archibald I.G., Bugg C.E., Carne A., Gover S., Helliwell J.R., Pickersgill R.W. and White S.W. (1983), EMBO JOURNAL 2, vol 6, p.1009-1014.
- Adams M.J., Archibald I.G., Helliwell J.R., Jenkins S.E. and White S.W. (1981) in "Structural studies on molecules of biological interest", (eds. G.Dodson, J.P. Glusker and D.Sayre), Clarendon Press, Oxford, pp. 328.
- Adams M.J., Helliwell J.R. and Bugg C.E. (1977), J. Mol. Biol. 112, 183-197.
- Allen M.B., (1953), Bacteriol. Rev. 17, 125-173.
- Allinson N.M. (1982), Nucl. Instrum. and Meth. 201, 53.
- Amelunxen R.E., Noelken M. and Singleton R. (1970), Arch. Biochem. Biophys., 141, 447-455.
- Arndt U.W., Champness J.N., Phizacherley R.P. and Wonacott A.J. (1973), J. Appl. Cryst., 6, 457-463.
- Arndt U.W., Greenhough T.J., Helliwell J.R., Howard J.A.K., Rule S.A. and Thompson A.W. (1982), Nature 298, 835-838.
- Arndt U.W. and Wonacott A.J. (1977), "The Rotation Method in Crystallography", North Holland, Amsterdam, New York, Oxford.
- Atkinson A., Phillips B.W., Callow D.S., Stoner W.R. and Bradford P.A. (1972), Biochem., J. 127, 63.
- Bartunik H.A., Clout P.N. and Robrahn B. (1981), J. Appl. Cryst. 14, 134-6.
- Bartunik H.A., Fourme R and Phillips J.C. (1982), in "Uses of Synchrotron Radiation in Biology", edited by Stuhrmann H.B., London, Academic Press.
- Betts S.A. and Mayer R.J. (1975), Biochem. J. 151, 263-270.
- Biesecker G., Harris J.I., Thierry J.C., Walker J.E. and Wonacott A.J., Nature 266, 328-333.
- Bigelow C.C. (1967), J. Theor. Biol. 16, 187-211.

Blake and Phillips (1962), in "Biological Effects of Ionising Radiation at the Molecular Level", IAEA symposium, p.183.

Blundell T.L. and Johnson L.N. (1976), "Protein Crystallography", London: Academic Press.

Bordas J. (1982), report to the instrumentation subgroup of ESP, Ad-hoc committee on synchrotron radiation.

Brandts J.F. (1967), in "Thermobiology", Rose A.H. (ed.), Academic Press Inc., New York.

Bridges R.B., Palumbo M.P. and Wittenberger C.L. (1975), J. Biol. Chem. 250, 6093-6100.

Brock T.D. (1967), Science 169, 1316-1318.

Compton A.H. and Allison S.K. (1935), "X-rays in theory and experiment", Van Nostrand, New York.

Crowther (1971), in "Molecular Replacement Method", ed. Rossmann M.G., (see reference).

Cutfield J.F., Cutfield S.M., Dodson E.J., Dodson G.G. and Sabesan M.N. (1974), J. Mol. Biol. 87, 23-30.

Debye P (1914), Annl. Phys. 43, 49.

Dodson E., Harding M.M., Hodgkin D.C. and Rossmann M.G. (1966), J. Mol. Biol. 16, 227.

Dodson G.G., Liddington R.C. and Vallely D.G. (1983), 1982-1983 Daresbury Laboratory Annual Report.

Dyson J.E.D., D'Orazio R.E. and Hanson W.H. (1973), Arch. Biochem. Biophys., 154, 623-635.

Eklund H., Nordstrom B., Zeppezauer E., Soderlund G., Boiwe T and Branden C.I. (1974), FEBS letters 44, 200.

Emdin, Peterson, Coulter, Ostberg, Falkner and Steiner (1973), 9th Internat. Cong. Biochem., Stockholm, 1-7 July 1973.

Ewald P.P. (1921), Z. Kristolleg. Miner. 56, 129.

Fisher H.F. (1964), Proceedings. Nat. Acad. Science., 51, 1285-1291.

Fontana A., Mantovanelli L., Boccu E and Veronese F.M., Int. J. Peptide Protein Res., 9, 329, (1977).

Fox and Holmes (1966), Acta. Cryst., 20, 886.

Golde A. and Warren G. (1980), Nucl. Instrum. and Meth., 172, 375-378.

Goldsach D.E. (1970), Biopolymers, 9, 247-252.

Greenhough T.J., Helliwell J.R. and Rule S.A. (1983), J. Appl. Cryst., 16, 242-250.

Greenhough T.J. and Machin P. (1983), Daresbury Laboratory technical memorandum. DL/CSD/TM23

Hachimori A., Muramatsu N and Nosoh Y. (1970), Biochim. Biophys. Act., 206, 426-437.

Harris J.I. (1975), in "Enzymes + Proteins from Thermophilic Micro-organisms", Zuber, p.416.

Helliwell J.R., Greenough T.J., Carr P.D., Rule S.A., Moore P.R., Thompson A.W and Worgan J.S. (1982), J. Phys. E. Sci., Instrum., 15, 1363-1372.

Helliwell J.R. and Fourme R (1983), "The ESRP as a facility for Protein Crystallography - A report and design study", submitted to the ESRP instrumentation sub-group committee 2/6/83.

Helliwell J.R. (1977), D. Phil. Thesis, Oxford.

Helliwell J.R. and Thompson A.W. (1983), Draft for Daresbury Laboratory Technical Memorandum.

Hendrix J., Koch M.H.J. and Bordas J. (1979), J. Appl. Cryst. 12, 467-472.

Hill E., Tsernoglou D., Webb L. and Banaszak L.J. (1972), J. Mol. Biol., 72, p.577.

Holmes and Blow (1966), "The Use of X-ray Diffraction in the Study of Protein and Nucleic Acid Structure", Interscience, New York.

Howard R.L. and Becker R.R. (1970), J. Biol. Chem., 245, 3186-3194.

International Tables for Crystallography (1968), General Editor K.Lonsdale, Kynock Press, Birmingham, England.

Kahn R., Fourme R., Gadet A., Javin J. Dumas C. and Andre D. (1982), J. Appl. Cryst., 15, 330.

Koffler H. and Yarbrough (1972), Conf. Extreme Environ., Ames Research Center, NASA, Moffet Field, Calif., 26-28 June (1972).

Larmor J. (1897), Phil. Mag., 44, 503.

Laue M.V. (1912), Sitz. Math. phys. Klasse bayer, Akad. Wiss., p.303.

Laue M.V. (1913), Ann. Physik, 41, 971.

Levy J., Rialdo G and Briltonen (1972), Biochemistry, 11, 4138-4144.

Lienard A. (1898), L'Edairage Elec., 16, 5.

Ljungdahl L., Brewer J.M., Neece S.H. and Fairwell T (1970), J. Biol. Chem., 245, 4791-4797.

Manning G.B., Campbell L.L. and Foster R.J. (1961), J. Biol. Chem., 236, 2958-2961.

Margoliash E. and Schejter A. (1966), Advanc. Protein Chem., 21, 113-286.

Materlik G. (1982), in "Uses of Synchrotron Radiation in Biology", edited by Stuhrmann H.B., London, Academic Press.

Mathews D.A., Alden R.A., Bolin J.T., Filman D.J., Freer S.T., Hamlin R., Hol W.G.J., Kislink R.L., Pastore E.J., Plante L.T., Xuong N.H. and Kraut J.J. (1978), J. Biol. Chem., 253, 6946.

Mathews D.A., Alden R.A., Bolin J.T., Freer S.T., Xuong N.H., Kraut J., Poe M., Williams M and Hoogsteen K. (1977), Science, N.Y., 197, 452.

Mathews B.W. and Remington S.J. (1974), Proceedings. Nat. Acad. Science., USA, 71, 4178.

Muramatsu N. and Nosoh Y. (1971), Arch. Biochem. Biophys., 144, 245-252.

Nixon P. (1973), D. Phil. Thesis, "A Crystallographic Comparison of Protein Molecules with Related Conformations", Wolfson College, Oxford.

Nixon P.E. and North A.C.T. (1976), Acta. Cryst., A32, 320.

Ohta Y., Ogura Y and Wada A. (1966), J. Biol. Chem., 241, 5919-5925.

Pearse B.M.F. and Harris J.I., (1973), FEBS Letters, 38, 49-52.

Pearse B.M.F. and Rosemeyer M.A. (1974), Eur. J. Biochem., 42, 213-223.

Perutz M.F. and Raidt H. (1975), Nature, 255, 256-259.

- Pfeuller S.L. and Elliot W.H. (1969), J. Biol. Chem. 244, 48-54.
- Popjak G. (1970), in "The Enzymes", (ed. P.D. Boyer), vol. II, p.116-215, New York, Academic Press.
- Procsal D. and Holton D. (1972), Biochemistry, 11, 1310.
- Rao S.N., Jyh-Hwang J. and Hartsuck J.A., Acta. Cryst. A36, 878-884.
- Rippa M., Signorini M and Pontremoli S. (1972), Arch. Biochem. Biophys., 150, 503-510.
- Rippa M., Signorini M. and Pontremoli S. (1970), Ital. J. Biochem., 19, p.361.
- Rippa M., Signorini M. and Pontremoli S. (1967), Eur. J. Biochem., 1, 170-178.
- Rossmann M.G. (1972), "Molecular Replacement Method", Gordon and Breach, New York.
- Rossmann M.G. and Blow D.M. (1962), Acta. Cryst., 15, 24-31.
- Rossmann M.G., Blow D.M., Harding M.M. and Collier E. (1964), Acta. Cryst., 17, 338.
- Rossmann M.G., Liljas A., Branden C.I. and Banaszak L.J. (1975), in "The Enzymes", 3rd edn. (ed. P.D. Boyer) vol. XI, p.61, Academic Press, New York.
- Ru-chang B., Cutfield S.M., Dodson E.J., Dodson G.G., Giordano F., Reynolds C.D. and Tolley S. (1983), Acta. Cryst., B39, 90-98.
- Scherga H.A. (1963), in "The Proteins, vol 1", Neurath H. (Ed.), Academic Press Inc. New York.
- Schulz G.E., Schirmer R.H., Sachsenheimer W. and Pai E.F. (1978), Nature Lond., 273, 120.
- Schwager P., Bartels K and Jones A. (1975), J. Appl. Cryst., 8, 275.
- Schwinger (1954), Proceedings. Nat. Acad. Sci., USA., 40, 132.
- Schwinger (1949), Phys. Rev., 75, 1912.
- Schwinger (1946), Phys. Rev., 70, 798.

Scott W.A. and Abramsky T. (1973), J. Biol. Chem., 248, 3535-3541.

Silverberg M. (1973)., D. Phil., Oxford.

Silverberg M. and Dalziel K. (1973), Eur. J. Biochem., 38, 229-238.

Singleton R. and Amelunxen R.E. (1973), Bacteriol. Rev., p.320-340.

Singleton R., Kimmel J.R. and Amelunxen R.E. (1969), J. Biol. Chem., 244, 1623-1630.

Smith J.L., Hendrickson W.A. and Addison A.W. (1983), Nature, 303, 86-88.

Sternberg M.J.E. and Thornton J.M. (1976), J. Mol. Biol., 705, 367.

Stuhrmann (1978), Quarterly review of Biophysics, 11, (1), 71-98.

Sugimoto S. and Nosoh Y. (1971), Biochim. Biophys. Acta., 235, 210-221.

Tanaka M., Haniu M., Matsueda G., Yasunobu K., Himes R.H., Akagi J.M., Barnes E.M. and Devanathan T. (1971), J. Biol. Chem., 246, 3953-3960.

Tollin P., Main P and Rossman M.G. (1966), Acta. Cryst., 20, 404-407.

Tollin P and Rossman M.G. (1966), Acta. Cryst., 21, 872.

Veronese F.M., Bocco E. and Fontana A. (1982), in "Methods in Enzymology", 90E, (Ed. W.A. Wood), Academic Press, New York.

Veronese F.M., Boccu E. and Fontana A. (1976), Biochemistry, 15, 4026.

Veronese F.M., Boccu E., Fontana A., Benassi C.A. and Scoffone E. (1974), Biochim., Biophys., Acta., 334, 31.

Villet R.H. and Dalziel K. (1972), Eur. J. Biochem., 27, 251-258.

Waugh W.F. (1954), *Advan. Protein Chem.*, 9, 326-437.

White J.L., Hackert M.L., Buchner M., Adams M.J., Ford G.C., Lentz P.J. Jr., Smiley I.E., Steindel S.J. and Rossman M.G. (1976), *J. Mol. Biol.*, 102, 759.

Wierenga R.K., de Jong R.J., Kalk K.H., Hol W.G.J. and Drenth J. (1979), *J. Mol. Biol.*, 131, p.55.

Wilson K.S., Stura E.A., Wild D.L., Todd R.J., Stuart D.I., Babu S., Jenkins J.A., Standing T.S., Johnson L.N., Fourme R., Kahn R., Gadet A., Bartels K.S. and Bartunik H.K. (1983), *J. Appl. Cryst.*

Wilson K. and Yeates D. (1979), *Acta. Cryst.* A35, 146-157.

Winick H. and Bienenstock A. (1978), *Ann. Rev. Nucl. Part. Sci.*, 28, 33.

Wonacott A.J. (1977), "Geometry of the Rotation Method", Chapter 7 of Ardnt U.W. and Wonacott A.J. (1977), - see reference.

Wonacott A.J. (1980), Unpublished program documentation for MOSCO.

Zappe H., Krohne-Ehrich G. and Schultz G.E. (1977), *J. Mol. Biol.*, 113, 141.

Zuber H. (Ed.), "Enzymes and Proteins from Thermophilic Micro-organisms", Birkhauser Verlag, Basel, (1975).

APPENDIX I Microfiche of stepper Motor Software

MARK I & MARK II STEPPER MOTOR SOFTWARE
ANDY.FORT

1of2

3RD PARTY COPYRIGHT
MATERIAL EXCLUDED
FROM DIGITISED THESIS

PLEASE REFER TO THE
ORIGINAL TEXT TO SEE
THIS MATERIAL